

OM of: US-09-877-665-4 to: Issued\_Patents\_NA.\* out\_format : pfs  
Date: Sep 15, 2002 1:30 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O/cg2n2.1/USPTO\_SPOOL/US09877665/runat\_13092002.083913.15672/app-query.fasta.1.103  
-B-issued\_Patents\_NA -QFMT=fastap -SUFFIX=p2n.rml  
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOEPC=0.000  
-LOEPC=0.000 -OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000  
-XGAPEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000  
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRIX=plowsum62 -TRANS=human40.cdi -LIST=45 -DOCALLIGN=200  
-THR.SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pfs -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09877665 @CGN1\_1.43 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMOUT=120 -WARN\_TIMOUT=30 -NO\_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-09-877-665-4  
Query length: 47  
Database: Issued\_Patents\_NA.\*  
Database sequences: 38353  
Database length: 122816752  
Search time (sec): 46.300000

score\_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation	
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-899-437-22			277.00	673.67	7.8e-30	2091
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-09-126-121-22			277.00	673.67	7.8e-30	2091
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-899-437-5			277.00	671.85	9.8e-30	2502
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-09-126-121-5			277.00	671.85	9.8e-30	2502
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-899-437-1			277.00	671.71	1.0e-29	2538
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-09-126-121-1			277.00	671.71	1.0e-29	2538
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-238-182-2			113.50	282.95	4.5e-08	192
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-238-182-2			113.50	282.95	4.5e-08	192
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-469-559-155			113.50	282.95	4.5e-08	192
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-469-559-155			113.50	282.95	4.5e-08	192
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-249-322A-155			113.50	282.95	4.5e-08	192
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-249-322A-155			113.50	282.95	4.5e-08	192
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-469-526A-155			113.50	282.95	4.5e-08	192
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-469-526A-155			113.50	282.95	4.5e-08	192
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-734-591A-155			113.50	282.95	4.5e-08	192
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-734-591A-155			113.50	282.95	4.5e-08	192
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-469-660-155			113.50	282.95	4.5e-08	192
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-469-660-155			113.50	282.95	4.5e-08	192
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-470-335-155			113.50	282.95	4.5e-08	192
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-470-335-155			113.50	282.95	4.5e-08	192
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-735-021-155			113.50	282.95	4.5e-08	192
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-735-021-155			113.50	282.95	4.5e-08	192
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-469-559-159			113.50	280.20	6.4e-08	252
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-469-559-159			113.50	280.20	6.4e-08	252
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-469-526A-159			113.50	280.20	6.4e-08	252
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-469-526A-159			113.50	280.20	6.4e-08	252
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-734-591A-159			113.50	280.20	6.4e-08	252
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-734-591A-159			113.50	280.20	6.4e-08	252
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-469-660-159			113.50	280.20	6.4e-08	252
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-469-660-159			113.50	280.20	6.4e-08	252
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-470-335-159			113.50	280.20	6.4e-08	252
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-470-335-159			113.50	280.20	6.4e-08	252
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-735-021-159			113.50	280.20	6.4e-08	252
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-735-021-159			113.50	280.20	6.4e-08	252
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-469-555B-158			113.50	279.62	6.9e-08	267
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-469-555B-158			113.50	279.62	6.9e-08	267
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-249-322A-158			113.50	279.62	6.9e-08	267
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-249-322A-158			113.50	279.62	6.9e-08	267
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-469-526A-158			113.50	279.62	6.9e-08	267
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-469-526A-158			113.50	279.62	6.9e-08	267
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-734-591A-158			113.50	279.62	6.9e-08	267
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-734-591A-158			113.50	279.62	6.9e-08	267
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-469-660-158			113.50	279.62	6.9e-08	267
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-469-660-158			113.50	279.62	6.9e-08	267

/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-470-335-158 + 113.50 279.62 6.9e-08 267  
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-470-335-158 + 113.50 279.62 6.9e-08 267  
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-734-591A-158 + 113.50 279.62 6.9e-08 267  
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-734-591A-158 + 113.50 279.62 6.9e-08 267  
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-469-555B-158 + 113.50 279.62 6.9e-08 267  
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-469-555B-158 + 113.50 279.62 6.9e-08 267  
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-249-322A-158 + 113.50 279.62 6.9e-08 267  
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-249-322A-158 + 113.50 279.62 6.9e-08 267  
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-469-526A-158 + 113.50 279.62 6.9e-08 267  
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-469-526A-158 + 113.50 279.62 6.9e-08 267  
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-734-591A-158 + 113.50 279.62 6.9e-08 267  
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-734-591A-158 + 113.50 279.62 6.9e-08 267  
/cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-469-660-158 + 113.50 279.62 6.9e-08 267  
/cg2n2.6/prodata/1/ina/6B.COMB.seq:US-08-469-660-158 + 113.50 279.62 6.9e-08 267

```

seq_name: /cg2n2.6/prodata/1/ina/6A.COMB.seq:US-08-899-437-22
seq_documentation_block:
; Sequence 22, Application US/08899437
; Patent No. 6121415
;
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Mary, Melanie Rose, Zhang, Dong Xiao
; TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 Inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatlin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,437
; FILING DATE: 24-Jul-1997
;
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Conley, Delire L.
; REGISTRATION NUMBER: 36,487
; REFERENCE/DOCKET NUMBER: P1084R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-2066
; TELEFAX: 650/952-9881
;
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2091 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
;
; FEATURE:
; NAME/KEY: Human NRG3B2 (hNRG3B2)
; LOCATION: 1-2091
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
;
; US-08-899-437-22
;
; alignment_scores:
; Quality: 277.00 Length: 47
; Ratio: 5.894 Gaps: 0
; Percent Similarity: 100.000 Percent Identity: 100.000
;
; alignment_block:
; US-09-877-665-4 x US-08-899-437-22 ..
;
; Align seg 1/1 to: US-08-899-437-22 from: 1 to: 2091
;
; 1 HsPhelysProCysArgAspLysAspLeuAlaTyrCysLeuAsnAsp1 17
; |||||
; 856 CACTTAACACCTCGCCGACACAGGACCTTCGATCTGCTCAATGATG 905
; |||||
; 17 yslucysphevalillegltThrleuThrglySerHslySHsCysArgC 34
; |||||
; 906 CAGAGCTTGTGATCGAAGAACCTCGACGATCCATTAACACTGCGGT 955
; |||||
; 34 yslucysgluclytyrGlnGlyValAlaGcysAspGlnPheLeu 47
; |||||

```

956 GCAAGAGGCTACCAAGAGGCTTGTGATCAATTTCTG 996

seq\_name: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:US-09-126-121-22

seq\_documentation\_block:

Sequence 22, Application US/09126121

Patent No. 6252051

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanle Rose, Zhang, Dong Xiao

TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/126,121

FILING DATE: 30-Jul-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Conley, Delidre L.

REGISTRATION NUMBER: 36,487

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-2066

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 2091 base pairs

TYPE: Nucleic Acid

STRANDEDNESS: Single

TOPOLOGY: Linear

FEATURE:

NAME/KEY: Human NRG3B2 (hNRG3B2)

LOCATION: 1-2091

IDENTIFICATION METHOD:

OTHER INFORMATION:

US-09-126-121-22

alignment\_scores:

Quality: 277.00

Ratio: 5.894

Percent Similarity: 100.000

Percent Identity: 100.000

Length: 47

Gaps: 0

Alignment block:

US-09-877-665-4 x US-09-126-121-22

Align seg 1/1 to: US-09-126-121-22 from: 1 to: 2091

1 HisphelyProCysArgAspLysAspLeuAlaTyrCysLeuSnaSpG1 17  
|||||  
856 CACTTCAAAACCTGCGGAGACAGACCTTGCATCTGCTCATGATGG 905  
|||||  
17 yglucysphevalileglutThrLeuThrglySerHisLysHisCysArgC 34  
|||||  
906 CGAGTCTTGTGTGATCGAAACCTGACCGGATCCCATAAACACTGTCGT 955  
|||||  
34 yslvsgluclytyrGlnGlyValArgCysAspGlnPheLeu 47  
|||||  
956 GCAAGAGAGGCTACCAAGAGGCTTGTGATCAATTTCTG 996

seq\_name: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:US-08-899-437-5

seq\_documentation\_block:

Sequence 5, Application US/08899437

Patent No. 6121415

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanle Rose, Zhang, Dong Xiao

TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/899,437

FILING DATE: 24-Jul-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Conley, Delidre L.

REGISTRATION NUMBER: 36,487

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-2066

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 2502 base pairs

TYPE: Nucleic Acid

STRANDEDNESS: Single

TOPOLOGY: Linear

FEATURE:

NAME/KEY: Human NRG3B1 (hNRG3B1)/nucleic acid seq.

LOCATION: 1-2502

IDENTIFICATION METHOD:

OTHER INFORMATION:

US-08-899-437-5

alignment\_scores:

Quality: 277.00

Ratio: 5.894

Percent Similarity: 100.000

Percent Identity: 100.000

Length: 47

Gaps: 0

Alignment block:

US-09-877-665-4 x US-08-899-437-5

Align seg 1/1 to: US-08-899-437-5 from: 1 to: 2502

1 HisphelyProCysArgAspLysAspLeuAlaTyrCysLeuSnaSpG1 17  
|||||  
999 CACTTCAAAACCTGCGGAGACAGACCTTGCATCTGCTCATGATGG 1048  
|||||  
17 yglucysphevalileglutThrLeuThrglySerHisLysHisCysArgC 34  
|||||  
1049 CGAGTCTTGTGATCGAAACCTGACCGGATCCCATAAACACTGTCGT 1098  
|||||  
34 yslvsgluclytyrGlnGlyValArgCysAspGlnPheLeu 47  
|||||  
1099 GCAAGAGAGGCTACCAAGAGGCTTGTGATCAATTTCTG 1139

seq\_name: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:US-09-126-121-5

seq\_documentation\_block:

Sequence 5, Application US/09126121

Patent No. 6252051

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
 TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genentech, Inc.  
 STREET: 1 DNA Way  
 CITY: South San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94080  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Winpatin (Genentech)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/126,121  
 FILING DATE: 30-Jul-1998  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Conley, Delidre L.  
 REGISTRATION NUMBER: 36,487  
 REFERENCE/DOCKET NUMBER: P1084R1D1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650/225-2066  
 TELEFAX: 650/952-9881  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2502 base pairs  
 TYPE: Nucleic Acid  
 STRANDEDNESS: Single  
 TOPOLOGY: Linear  
 FEATURE:  
 NAME/KEY: Human NR3B1(hNR3B1)/nucleic acid seq.  
 LOCATION: 1-2502  
 IDENTIFICATION METHOD:  
 OTHER INFORMATION:  
 US-09-126-121-5  
 alignment\_scores:  
 Quality: 277.00 Length: 47  
 Ratio: 5.894 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000  
 alignment\_block:  
 US-09-877-665-4 x US-09-126-121-5 ..  
 Align seg 1/1 to: US-09-126-121-5 from: 1 to: 2502  
 1 HsPhelysProCysArgAspLysAspLeuAlaTyrCysLeuAsnAspG1 17  
 |||||||  
 999 CACTCAAAACCTGCGGAGACAAGACCTTGCATCTACTGCTCAATGATG 1048  
 |||||||  
 17 yGlucyspheValIleGluThrLeuThrgLysSerHisLysHisCysArgC 34  
 |||||||  
 1049 CGAGGCTTTGGATGCAAGACCTGACCGGATCCATAAACACTGCGGT 1098  
 |||||||  
 34 yLysGluGluTyrGlnGlyValArgCysAspGlnPheLeu 47  
 |||||||  
 1099 GCAGAGAGAGCTACCAAGAGATCCCTGTGTGATCAATTCTGTG 1139  
 |||||||  
 seq\_name: /cgn2\_6/ptodata/1/lna/6A\_COMB.seq:US-08-899-437-1  
 seq\_documentation\_block:  
 : Sequence 1, Application US/08899437  
 : Patent No. 6121415  
 : GENERAL INFORMATION:  
 : APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
 : TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
 : NUMBER OF SEQUENCES: 23  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Genentech, Inc.  
 : STREET: 1 DNA Way  
 : CITY: South San Francisco

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genentech, Inc.  
 STREET: 1 DNA Way  
 CITY: South San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94080  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Winpatin (Genentech)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/899,437  
 FILING DATE: 24-Jul-1997  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Conley, Delidre L.  
 REGISTRATION NUMBER: 36,487  
 REFERENCE/DOCKET NUMBER: P1084R1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650/225-2066  
 TELEFAX: 650/952-9881  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2538 base pairs  
 TYPE: Nucleic Acid  
 STRANDEDNESS: Single  
 TOPOLOGY: Linear  
 FEATURE:  
 NAME/KEY: mouse NR3 nucleic acid  
 LOCATION: 1-2538  
 IDENTIFICATION METHOD:  
 OTHER INFORMATION:  
 US-08-899-437-1  
 alignment\_scores:  
 Quality: 277.00 Length: 47  
 Ratio: 5.894 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000  
 alignment\_block:  
 US-09-877-665-4 x US-08-899-437-1 ..  
 Align seg 1/1 to: US-08-899-437-1 from: 1 to: 2538  
 1 HsPhelysProCysArgAspLysAspLeuAlaTyrCysLeuAsnAspG1 17  
 |||||||  
 1150 CACTCAAAACCTGCGGAGACAAGACCTGCGGTATGCTCAATGATG 1199  
 |||||||  
 17 yGlucyspheValIleGluThrLeuThrgLysSerHisLysHisCysArgC 34  
 |||||||  
 1200 TGAATGCTTTGGATGAGACCTGACAGGATCCATAAGCACTGCTGT 1249  
 |||||||  
 34 yLysGluGluTyrGlnGlyValArgCysAspGlnPheLeu 47  
 |||||||  
 1250 GCAGAGAGAGCTACCAAGAGATCCCTGTGTGATCAATTCTGTG 1290  
 |||||||  
 seq\_name: /cgn2\_6/ptodata/1/lna/6B\_COMB.seq:US-09-126-121-1  
 seq\_documentation\_block:  
 : Sequence 1, Application US/09126121  
 : Patent No. 6252051  
 : GENERAL INFORMATION:  
 : APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
 : TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
 : NUMBER OF SEQUENCES: 23  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Genentech, Inc.  
 : STREET: 1 DNA Way  
 : CITY: South San Francisco

```

STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/126,121
FILING DATE: 30-JUL-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Delidre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2538 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
FEATURE:
NAME/KEY: mouse NR3 nucleic acid
LOCATION: 1-2538
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-09-126-121-1

```

```

alignment_scores:
Quality: 277.00      Length: 47
Ratio: 5.894        Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000

```

```

alignment_block:
US-09-877-665-4 x US-09-126-121-1 ..

```

```

Align seg 1/1 to: US-09-126-121-1 from: 1 to: 2538

```

```

1 HisphelysProcysaArgAspLysAspLeuAlaTyrCysLeuAsnAspG1 17
|||||
1150 CACTTCAACCTGTGCGAGACAGACCTGGCGTATGTCTCAATGATGG 1199
17 yGIuCySpheValIleGluThrLeuThrglySerHisLysHisCysArgC 34
|||||
1200 TGAATGCTTGTGATTGAGACCTGACAGGATCCCATAGACACTCTCGT 1249
34 yALySGlUGlYTYrGInGlyValAlaGcYsAspGlnPhelau 47
|||||
1250 GCAAGGAAGGCTACCAAGAGAGCTGTGTGATCAATTTCTG 1290

```

```

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-09-238-182-2

```

```

seq documentation block:
; Sequence 2, Application US/09238182
; Patent No. 6080845
; GENERAL INFORMATION:
; APPLICANT: Carnahan, Josette F.
; TITLE OF INVENTION: MONOCLONAL ANTIBODY AGAINST UTRICULAR EPITHELIUM
; FILE REFERENCE: A-532A
; CURRENT APPLICATION NUMBER: US/09/238,182
; CURRENT FILING DATE: 1999-01-28
; EARLIER APPLICATION NUMBER: 09/129,549
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 156
; TYPE: DNA

```

```

; ORGANISM: Human
US-09-238-182-2

```

```

alignment_scores:
Quality: 116.50      Length: 48
Ratio: 3.149        Gaps: 1
Percent Similarity: 77.083  Percent Identity: 33.333

```

```

alignment_block:
US-09-877-665-4 x US-09-238-182-2 ..

```

```

Align seg 1/1 to: US-09-238-182-2 from: 1 to: 156

```

```

1 HisphelysProcysaArgAspLysAspLeuAlaTyrCysLeuAsnAspG1 17
|||||
4 CACTTGTAAATGTGCGGAGAGAGAAACTTCTGTGTGATGAGAGG 53
17 yGIuCySpheValIleGluThrLeuThrglySerHisLysHis...CysA 33
|||||
54 GAGAGCTCTCATGTGTAAGACCTTCAACCCCTGAGATACTGTGCA 103
33 iGcYsLysGluGlyTYrGInGlyValAlaGcYsAspGlnPhelau 47
|||||
104 AGTCCCAACCTGATGACTGAGAGCAAGATGTCAAACTACGTA 147

```

```

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-036-555B-155

```

```

seq documentation block:
; Sequence 155, Application US/08036555B
; Patent No. 5530109
; GENERAL INFORMATION:
; APPLICANT: Goodheart, Andrew; Stroobant, Paul;
; APPLICANT: Minchetti, Luisa; Waterfield, Michael; Marchioni, Mark;
; APPLICANT: Chen, Mao Su; Hiles, Ian
; TITLE OF INVENTION: Glial Mitogenic Factors, Their
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Feltz & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/036,555B
; FILING DATE: 24-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APRIL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.K. 91 07566.3
; FILING DATE: 10-APRIL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: LDD 5250.4

```



```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/863,703
FILING DATE: 03-APRIL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.K. 91 07566.3
FILING DATE: 10-APRIL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Tsai, Christine H.
REGISTRATION NUMBER: 34,266
REFERENCE/DOCKET NUMBER: LUD 5250.4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 155:
SEQUENCE CHARACTERISTICS:
LENGTH: 192
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-469-569-155

Alignment_scores:
Quality: 113.50 Length: 46
Ratio: 3.338 Gaps: 1
Percent Similarity: 73.913 Percent Identity: 34.783

Alignment_block:
US-09-877-665-4 x US-08-469-569-155 ..

Align seg 1/1 to: US-08-469-569-155 from: 1 to: 192

1 HisphelysProCysArgAspIleuAlaTyrCysLeuAsnAspG1 17
||||:||||:||||:||||:||||:||||:||||:||||: 11
4 CATCTGTCAAGTGTGCAGAGAGAGAAAACCTTCTGTGTGAATGGAG 53
17 YGluCysPheValIleGluThrLeuThrGlySerHisLysHis...CysA 33
||||:||||:||||:||||:||||:||||:||||:||||: 11
54 CGAGCGCTTCAATGGTGAAGAAGACCTTCAATCCCTCAAGTACTGTGCA 103
33 rGcYsLysGluGlyTyrGlnGlyValArgCysAspGln 45
||||:||||:||||:||||:||||:||||:||||:||||: 11
104 AGTGCACAACCTGGATTCAGTCTGGAGCGAGATGACTGAG 141

seq_name: /cgr2_6/prodata/1/lna/5A_COMB.seq:US-08-249-322A-155
seq_documentation_block:
; Sequence 155, Application US/08249322A
; Patent No. 5716930
GENERAL INFORMATION:
APPLICANT: Goodheart, Andrew; Stroobant, Paul;
APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionti, Mark;
APPLICANT: Chen, Mao Su; Hiles, Ian
TITLE OF INVENTION: Glial Mitogenic Factors, Their
TITLE OF INVENTION: Preparation and use
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felte & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/249,322A
FILING DATE: 26-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

```

```
seq_name: /cgn2-6/ptodata/1/1na/5B-COMB.seq:us-08-734-591A-155
seq_documentation_block:
  Sequence 155, Application US/08734591A
  Patent No. 5854220
  GENERAL INFORMATION:
  APPLICANT: Goodearl, Andrew
  APPLICANT: Stroobant, Paul
```

```

APPLICANT: Minghetti, Luisa
APPLICANT: Waterfield, Michael
APPLICANT: Hiles, Ian
APPLICANT: Matchonni, Mark
APPLICANT: Chen, Mario
TITLE OF INVENTION: GIAL MITOGENIC FACTORS, THEIR
TITLE OF INVENTION: PREPARATION AND USE
NUMBER OF SEQUENCES: 187
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Ebling, LLP
STREET: 176 Federal Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible Pentium
OPERATING SYSTEM: Windows95
SOFTWARE: Wordperfect (Version 7.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,591A
FILING DATE: 22-OCT-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/470,335
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/036,555
FILING DATE: 03-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,173
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/940,389
FILING DATE: 03-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/907,138
FILING DATE: 30-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/863,703
FILING DATE: 03-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 91 07566.3
FILING DATE: 10-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bleker-Brady, Kristina
REGISTRATION NUMBER: 39,109
REFERENCE/DOCKET NUMBER: 04585/00200P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 428-0200
TELEFAX: (617) 428-7045
TELEX:

INFORMATION FOR SRO ID NO: 155:
SEQUENCE CHARACTERISTICS:
LENGTH: 192
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-734-591A-155

alignment_scores:
Quality: 113.50 Length: 46
Ratio: 3.338 Gaps: 1
Percent Similarity: 73.913 Percent Identity: 34.783

alignment_block:
US-09-877-665-4 x US-08-734-591A-155 ..
Align seg 1/1 to: US-08-734-591A-155 from: 1 to: 192

1 HspHbelyProCysArGspIysAspLeuAlaTyrCysLeuAsnAspI 17

```

```

17 yglucyshevaliiegutnrlenthglyserHisylshts...Cysa 33
54 CGAGGCTTCATGGGAGAAAGACCTTCAAAATCCTCAAGATACTTGGCA 103
104 AGTCCCACTCGATTCCTGAGCGAGATCTACTGAG 141
33 rgcylsyagluiglytrfclngllyvalarqcyasapgin 45
seq_name: /cgn2.6/prodata/1/lna/5b_comb.seq:US-08-469-660-155
seq_documentation_block:
  Sequence 155, Application US/08469660
  Patent No. 5876973
  GENERAL INFORMATION:
  APPLICANT: Gwynne, David I.; Marchionni, Mark;
  APPLICANT: McBurney, Robert N.
  TITLE OF INVENTION: INHIBITORS OF CELL PROLIFERATION,
  TITLE OF INVENTION: THEIR PREPARATION AND USE
  NUMBER OF SEQUENCES: 184
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Fish & Richardson
  STREET: 225 Franklin Street
  CITY: Boston
  STATE: Massachusetts
  ZIP: 0211-2804
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
  COMPUTER: IBM
  OPERATING SYSTEM: PC-DOS
  SOFTWARE: Wordperfect
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/469,660
  FILING DATE:
  CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/011,396
  FILING DATE: 29-JAN-1993
  PRIOR APPLICATION DATA: 07/984,085
  APPLICATION NUMBER: 01-DEC-1992
  FILING DATE: 01-DEC-1992
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 07/951,747
  FILING DATE: 25-SEP-1992
  PRIOR APPLICATION DATA: 07/927,337
  APPLICATION NUMBER: 10-AUG-1992
  FILING DATE: 10-AUG-1992
  ATTORNEY/AGENT INFORMATION:
  NAME: Clark, Paul T.
  REGISTRATION NUMBER: 30,162
  REFERENCE/DOCKET NUMBER: 04585/017004
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: (617) 542-5070
  TELEFAX: 200154
  INFORMATION FOR SEQ ID NO: 155:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 192
  TYPE: nucleic acid
  STRANDEDNESS: double
  TOPOLOGY: linear
US-08-469-660-155
alignment_scores:
  Quality: 113.50 Length: 46
  Ratio: 3.338 Gaps: 1
  Percent Similarity: 73.913 Percent Identity: 34.783
alignment_block:
US-08-877-665-4 x US-08-469-660-155

```

```

1 HisPheIyLPrOCcSvrtgspLySaSpLeuAlatYrCysLeuAaaSPGL 17
   ||||| :||| ::|||::|||::|||::|||::|||::|||::|||
4 CACCTTGTCAAAGTGTGTCACAGAAGGAAGAAAACCTTTCTGTGGAAATGGAG 53
   ||||| :||| ::|||::|||::|||::|||::|||::|||
17 yGIuCySPheaValleGIuThrLeuThrGISeHrlsHis... CysA 33
   ||||| :||| ::|||::|||::|||::|||::|||::|||
54 CGAGTCCCTCAISGTSITAAAAACCCTTCAAATCCCTCAAGAAATCCTTGACA 103
   ||||| :||| ::|||::|||::|||::|||::|||::|||
33 rGCysLySseluGIyTrGIInGIyValArGcYSaSPGLn 45
   ::|||:: :|||:: :|||::|||::|||::|||::|||::|||
104 AGTGCACACTCGATTCACTGSAGGCAATGTACTGAG 141

```

```
seq_documentation_block:
```

Patent No. 6087323

APPLICANT: Gwynne,

APPLICANT: Marchionni, Mark A.

APPLICANT: Goldin, Stanley M.

TITLE OF INVENTION: USE OF NEUREGULINS AS MODULATORS OF

FILE REFERENCE: 04585/041001

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

```
! SOFTWARE: FastSEQ for Windows Version 4.0
! SEQ ID NO 1
```

; LENGTH: 192  
; TYPE: DNA

ORGANISM: Homo sapiens

NAME/KEY: CDS  
LOCATION: (1)

05-08-341-018-61

alignment\_scores:

Ratio:	3.338
Ratio:	11.0:20

Percent Identity: 34.783

US-09-877-665-4 x US-08-341-018-61

Align seg 1/1 to: US-08-341-018-61 from: 1 to: 192

```

1 HisPheIysProCysArgAspLysSplenuLatYrCLeuAsnAspG1 17
4 CATCTTGCAGATGTGCGAAGAAGAAACTTCCTGTGCANTGGACG 53
17 YGIcLysPheValIleGluThrIleuThngLySerHisLysAlts...Cysa 33
54 CGAGTGCTTACGTGGAAAGACCTTTCAATACCCTCAAGATATCTTGGA 107
```

```

33  rgCysLysgluglytyrGlnIyValAlrgCysaspGln 45
:::|::: |:::| |:::| |:::| |:::| |:::| :::
104 AGTGCACACCTGGATTTCACTGGAGCGAGATGTACTGAG 141

```

seq\_name: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:US-08-470-335-155

```
seq_documentation_block:
```

; Sequence 155, Application US/08470335F  
Patent No. 5147100

Patent No. 6147190

: APPLICANT: GOODBYE  
: GENERAL INFORMATION:  
: GENERAL INFORMATION:

APPLICANT: GOODEARL, ANDREW

APPLICANT: SIKOUBANI, PAUL

;  
; APPLICANT: MINGHETTI, LUISA  
;  
; APPLICANT: WATERFIELD, MICHAEL  
;  
; APPLICANT: MARCHIONNI, MARK  
;  
; APPLICANT: CHEN, MARIO S.  
;  
; APPLICANT: HILES, IAN  
;

; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR  
; TITLE OF INVENTION: PREPARATION AND USE  
; FILE REFERENCE: 04505/000000

```

; FILE REFERENCE: 04585/00200B
; CURRENT APPLICATION NUMBER: US/08/470,335H
; CURRENT FILING DATE: 10/05/08

```

```

; CURRENT FILING DATE: 1995-06-06
;
; E# IER APPLICATION NUMBER: 08/036,555
; E# IER FILING DATE: 1993-03-24
;

```

EA. IER FILING DATE: 1993-03-24  
NUMBER OF SEQ ID NOS: 252  
SOFTWARE: FastSeq for Windows Version 1.0

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 155
; LENGTH: 192
```

LENGTH: 192  
TYPE: DNA  
ORGANISM: Homo sapiens

```

; ORGANISM: Homo sapiens
;
; FEATURE:
; NAME/KEY: CDS

```

```

; NAME/NEI: CDS
; LOCATION: (1)...(189)
ITS-08-470-335-155

```

05-06-4/U-333-155

alignment\_scores:

Quality:	113.50
Ratio:	3.338
Percent similarity:	73.913

alignment block: Percent Similarity: 73.913 Percent Identity: 34.783

```
assignment_block: US-09-877-665-4 x US-08-470-335-155
```

Align seg 1/1 to: US-08-470-335-155 from: 1 to: 192

```

1 HisPheLysProCysArgAspLysAspLeuAlaTyrCysLeuAsnAspG1 17
2  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4 CATCTTGCATAGTGTGCAGAGACAGACAAAATTCTGTGTGAAATGGAGC 53
17 YGLCysPheValIleGluThrIleuThrGlySerHisLysHis...CysA 33
18 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
54 CGATGCTCTTACGGTGAAGACCTTTGAAATCCCTCAAGATAACTTGTGCA 103

```

[illegible]





CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple  
 CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,  
 CC and Meniere's disease. They can also be used to treat neuropathies  
 CC associated with systemic disease including post-polio syndrome,  
 CC hereditary neuropathies including Charcot-Marie-Tooth disease,  
 CC Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's  
 CC disease, metachromatic leukodystrophy, Fabry's disease and  
 CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of  
 CC skeletal muscle, such as muscular dystrophy or diseases caused by  
 CC for detection, diagnosis, for the production of transgenic or  
 CC knockout animals or for drug screening.

XX Sequence 2091 BP; 528 A; 633 C; 515 G; 415 T; 0 other;

# Alignment\_scores:

Quality:	277.00	Length:	47
Ratio:	5.894	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

# Alignment\_block:

US-09-877-665-4 x AAX06989

Align seg 1/1 to: AAX06989 from: 1 to: 2091

```

1 HisphelyProCysArgAspLysAspLeuAlaTYrCysLeuAsnAspG1 17
856 CACTTCAAAACCTGCGGAGACAGACCTTGCATCTCTCATGATGG 905
17 YGlucysPheValIleGluThrLeuThyGlySerHisLysHisCysArgC 34
906 CGAGTGCCTTTGTGATCGAAACCTGACCGGATCCCATTAACACGTGCGT 955
34 ySLySGlUGlYTYrGlnGlyValAlrGcYsAspGlnPheLeu 47
956 GCAAGAAGAGGCTACCAAGAGAGTCCGTTGTGATCAATTTCTG 996

```

seq\_name: /SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1999.DAT: AAX36423

# seq\_documentation\_block:

ID AAX36423 standard; DNA; 2199 BP.

XX AAX36423;

DT 06-JUL-1999 (first entry)

DE Human heregulin-like factor coding sequence.

KW Human heregulin-like factor; HLF; cell growth regulator; diagnosis;  
 KW neural system disorder; cancer; ss.

OS Homo sapiens.

PN WO9857989-A1.

PD 23-DEC-1998.

PF 16-JUN-1998; 98WO-US12403.

PR 17-JUN-1997; 97US-0049942.

PA (HUMA-) HUMAN GENOME SCI INC.

PA (GEOU) UNIV GEORGETOWN.

PI H13421 MM, King CR, Ruben SM, Young P;

DR WPI; 1999-095327/08.

DR P-PSDB; AAY05451.

PT New isolated heregulin-like factor - used to develop products for  
 PT the diagnosis and treatment of disorders involving regulation of  
 PT cell growth, particularly cancers

XX Claim 2; Page 86-87; 118pp; English.

PS This sequence encodes the human heregulin-like factor (HLF) of the  
 CC invention. The HLF is involved in the regulation of cell growth.  
 CC Detection of different levels of expression of the HLF gene can be used  
 CC for the diagnosis of disorders, e.g. in the neural system. In  
 CC particular, detection of different levels of HLF gene expression in cells  
 CC or body fluid of an individual can be used for diagnosing cancer. The  
 CC products can also be used in the treatment of disorders involving  
 CC abnormal levels of HLF activity.

XX Sequence 2199 BP; 689 A; 445 C; 462 G; 603 T; 0 other;

# Alignment\_scores:

Quality:	277.00	Length:	47
Ratio:	5.894	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

# Alignment\_block:

US-09-877-665-4 x AAX36423

Align seg 1/1 to: AAX36423 from: 1 to: 2199

```

1 HisphelyProCysArgAspLysAspLeuAlaTYrCysLeuAsnAspG1 17
92 CACTTCAAAACCTGCGGAGACAGACCTTGCATCTCTCATGATGG 141
17 YGlucysPheValIleGluThrLeuThyGlySerHisLysHisCysArgC 34
142 CGAGTGCCTTTGTGATCGAAACCTGACCGGATCCCATTAACACGTGCGT 191
34 ySLySGlUGlYTYrGlnGlyValAlrGcYsAspGlnPheLeu 47
192 GCAAGAAGAGGCTACCAAGAGAGTCCGTTGTGATCAATTTCTG 232

```

seq\_name: /SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1999.DAT: AAX06988

# seq\_documentation\_block:

ID AAX06988 standard; CDNA; 2502 BP.

XX AAX06988;

DT 10-MAY-1999 (first entry)

DE Human heregulin related ligand NRG3 CDNA.

KW Neuregulin related ligand; NRG3; hNRG3B1; human; ErbB4 receptor;  
 KW signal transduction; nervous system disorder; neurodegeneration;  
 KW neuropathy; therapy; diagnosis; ss.

OS Homo sapiens.

FN Key Location/Qualifiers

FT CDS 144..2306

FT /\*tag= a

PN WO9902681-A1.

PD 21-JAN-1999.

PF 30-JUN-1998; 98WO-US13411.

PR 24-JUL-1997; 97US-0899437.

PR 09-JUL-1997; 97US-0052019.

PA (GETH) GENENTECH INC.

PI Godowski PJ, Mark MR, Zhang D;

DR WPI; 1999-120882/10.

DR P-PSDB; AAW97618.

XX New isolated neuroguilin related ligand-3 - used to develop products  
 PT for treating nervous system disorders, e.g. stroke, ischaemia,  
 PT infection, malignancy, Alzheimer's disease or Down's syndrome  
 XX  
 PS Example 1; Page 64-66; 101pp: English.

CC This cDNA clone encodes human neuroguilin related ligand NRG3 (see  
 CC AAW97618), a novel member of the epidermal growth factor (EGF)-like  
 CC family of protein ligands that binds to the ErbB4 receptor, but not  
 CC to the ErbB2 or ErbB3 receptor, and which activates ErbB4 receptor  
 CC tyrosine phosphorylation. A partial human NRG3 cDNA clone was  
 CC obtained from a foetal brain cDNA library using a probe (see  
 CC AAX06981) based on an isolated expressed sequence tag sequence (see  
 CC AAX06990). Additional 5' sequence was obtained by anchored PCR. A  
 CC nucleic acid comprising the human NRG3 open reading frame cloned in  
 CC an expression vector is designated PRK5.tk.neo.NRG3B1 (ATCC 209157).  
 CC An alternatively spliced form of human NRG3 nucleic acid, hNRG3B2,  
 CC lacking nucleotides 1585-1656, is given in AAX06989. The chromosomal  
 CC localisation of human NRG3 was mapped to 10q22. The invention  
 CC provides human and murine NRG3 polynucleotides (see also AAX06987) and  
 CC polypeptides (see also AAW97617), expression vectors, host cells and  
 CC methods for the recombinant production of NRG3s. The polynucleotides  
 CC and polypeptides can be used to enhance the survival, proliferation  
 CC or differentiation of cells having the ErbB4 receptor in vivo and in  
 CC vitro. They can be used to prevent or treat damage to a nerve or  
 CC damage to other NRG3-expressing or NRG3-responsive cells, e.g.  
 CC brain, heart, or kidney cells. In particular, they can be used to  
 CC treat diseases which involve neural cell growth such as demyelination,  
 CC or damage or loss of glial cells (e.g. multiple sclerosis). They can  
 CC be used to treat patients whose nervous system has been damaged by  
 CC e.g. trauma, surgery, stroke, ischaemia, infection, metabolic  
 CC disease, nutritional deficiency, malignancy, or toxic agents. NRG3  
 CC can also be used to treat motor neuron disorders such as amyotrophic  
 CC lateral sclerosis (Lou Gehrig's disease), Bell's palsy, and  
 CC conditions involving spinal muscular atrophy or paralysis. NRG3  
 CC can be used to treat human neurodegenerative disorders such as  
 CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple  
 CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,  
 CC and Meniere's disease. They can also be used to treat neuropathies  
 CC associated with systemic disease including post-polio syndrome,  
 CC hereditary neuropathies including Charcot-Marie-Tooth disease, Krabbe's  
 CC Refsum's disease, abetalipoproteinemia, Fabry's disease and  
 CC disease, metachromatic leukodystrophy, Fanger's disease and  
 CC Dejerine-Sottas syndrome, to treat disease or diseases caused by  
 CC skeletal or smooth muscle wasting. The products can also be used  
 CC for detection, diagnosis, for the production of transgenic or  
 CC knockout animals or for drug screening.

XX Sequence 2502 BP; 647 A; 723 C; 614 G; 517 T; 1 other;

# alignment\_scores:

Quality: 277.00 Length: 47  
 Ratio: 5.894 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

# alignment\_block:

us-09-877-665-4 x AAX06988

Align seg 1/1 to: AAX06988 from: 1 to: 2502

1 HisPhelysProCysArGAspLysAspLeuAlaIatYrCysLeuAsnAspGI 17  
 999 CACTTCAAACCCCTCCGAGACAGACAGACCTTCGCATACGTCTCAATGATG 1048  
 17 yGluCysPheValIleGIuThrIeuThrngIySerHisLysHisCysArGc 34  
 1049 CGAGTGTCTTGTGATCGAAACCCCTGACCCGATCCCATAAACACATGCGGT 1098  
 34 ySLysGluGlyTYrGluGlyValArgCysAspGlnPheLeu 47  
 ||||||||||||||||||||||||||||||||||||||||||||

1099 GCAAGAGAGCTACCAAGAGATCCGTTGTGATCAATTCTG 1139

seq\_name: /SID51/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: AAX06987

seq\_documentation\_block:

ID AAX06987 standard; CDNA: 2538 BP.

AC AAX06987;

DT 10-MAY-1999 (first entry)

DE Mouse neuroguilin related ligand NRG3 cDNA.

KW Neuroguilin related ligand; NRG3; mouse; ErbB4 receptor;

KW signal transduction; nervous system disorder; neurodegeneration;

KW neuropathy; therapy; diagnosis; ss.

XX Mus sp.

XX Key Location/Qualifiers

XX CDS 289..2430

XX FT /\*tag= a

XX PN M09902681-A1.

XX PF 30-JUN-1998; 98MO-US13411.

XX PR 24-JUL-1997; 97US-0899437.

XX PR 09-JUL-1997; 97US-0052019.

XX PA (GETH ) GENENTECH INC.

XX PI Godowski PJ, Mark MR, Zhang D;

XX DR NPI; 1999-120882/10.

XX DR P-PADB; AAW97617.

XX PT New isolated neuroguilin related ligand-3 - used to develop products

XX PT for treating nervous system disorders, e.g. stroke, ischaemia,

XX PT infection, malignancy, Alzheimer's disease or Down's syndrome

XX PS Example 1; Page 57-59; 101pp: English.

CC This cDNA clone encodes murine neuroguilin related ligand NRG3 (see  
 CC AAW97617), a novel member of the epidermal growth factor (EGF)-like  
 CC family of protein ligands that binds to the ErbB4 receptor, but not  
 CC to the ErbB2 or ErbB3 receptor, and which activates ErbB4 receptor  
 CC tyrosine phosphorylation. To clone murine NRG3 cDNA, degenerate  
 CC primers based on regions (see AAW97623-24) proximal to the  
 CC transmembrane domain of a partial human cDNA were used to screen  
 CC a mouse brain cDNA library. An isolated clone was used as a  
 CC probe to obtain 2 further, overlapping clones that provided the  
 CC entire open reading frame. A nucleic acid comprising the murine  
 CC NRG3 open reading frame cloned into an expression vector is  
 CC designated pLXSN.NRG3 (ATCC 209156). The invention provides human  
 CC and murine NRG3 polynucleotides (see also AAX06988) and polypeptides  
 CC (see also AAW97618), expression vectors, host cells and methods for  
 CC the recombinant production of novel NRG3s. The polynucleotides and  
 CC polypeptides can be used to enhance the survival, proliferation or  
 CC differentiation of cells having the ErbB4 receptor in vivo and in  
 CC vitro. They can be used to prevent or treat damage to a nerve or  
 CC damage to other NRG3-expressing or NRG3-responsive cells, e.g.  
 CC brain, heart, or kidney cells. In particular, they can be used to  
 CC treat diseases which involve neural cell growth such as demyelination,  
 CC or damage or loss of glial cells (e.g. multiple sclerosis). They can  
 CC be used to treat patients whose nervous system has been damaged by  
 CC e.g. trauma, surgery, stroke, ischaemia, infection, metabolic  
 CC disease, nutritional deficiency, malignancy, or toxic agents. NRG3  
 CC can also be used to treat motor neuron disorders such as amyotrophic  
 CC lateral sclerosis (Lou Gehrig's disease), Bell's palsy, and  
 CC conditions involving spinal muscular atrophy or paralysis. NRG3



CC can be used to treat human neurodegenerative disorders such as  
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple  
CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,  
CC and Meniere's disease. They can also be used to treat neuropathies  
CC associated with systemic disease including post-polio syndrome,  
CC hereditary neuropathies including Charcot-Marie-Tooth disease,  
CC Refsum's disease, abetalipoproteinemia, Ranganer disease, Krabbe's  
CC disease, metachromatic leukodystrophy, Fabry's disease and  
CC Dystonia-torax syndrome, to treat disease of skeletal muscle of  
CC smooth muscle, such as muscular dystrophy or diseases caused by  
CC skeletal or smooth muscle wasting. The products can also be used  
CC for detection, diagnosis, for the production of transgenic or  
CC knockout animals or for drug screening.  
XX  
SQ Sequence 2538 BP; 606 A; 785 C; 649 G; 498 T; 0 other;

alignment\_scores:  
Quality: 277.00 Length: 47  
Ratio: 5.894 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-877-665-4 x AAX06987 ..

Align seg 1/1 to: AAX06987 from: 1 to: 2538

1 HisphelyProCyArgAspLysAspLeuAlaTyrCysLeuAsnAspGcl 17  
|||  
1150 CACTTCACAACTGCTGAGACAAAGACCTGCGTATTCCTCAAGATG 1199  
|||  
17 yelysgluyltyrtingllyalaargcysaspglnphelu 47  
|||  
1200 TAAATGCTTTGTGATTGACACCTGACAGATCCATTAAGCACCTGCGT 1249  
|||  
34 yelysgluyltyrtingllyalaargcysaspglnphelu 47  
|||  
1250 GCAAGGAAGGCTACCAAGAGATCGTGTGATCAATTCTG 1290  
|||

seq\_name: /SIDSL/9cdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX36424

seq\_documentation\_block:  
ID AAX36424 standard; DNA: 536 BP.  
XX  
AC AAX36424:  
XX  
DT 06-JUL-1999 (first entry)  
XX  
DE Human heregulin-like factor coding sequence clone HAGFE38R.  
XX  
KW Human heregulin-like factor; HLF; cell growth regulator; diagnosis;  
KW neural system disorder; cancer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9857989-A1.  
XX  
PD 23-DEC-1998.  
XX  
PF 16-JUN-1998; 98WO-US12403.  
XX  
PR 17-JUN-1997; 97US-0049942.  
XX  
PA (HOMA-) HUMAN GENOME SCI INC.  
PA (GECU) UNIV GEORGETOWN.  
XX  
PI H1Jaz1 NM, King CR, Ruben SM, Young P;  
XX  
WP1: 1999-095327/08.  
XX  
PT New isolated heregulin-like factor - used to develop products for  
PT the diagnosis and treatment of disorders involving regulation of  
PT cell growth, particularly cancers

XX  
PS Claim 20; Page 90-91; 118pp; English.  
XX  
CC This sequence encodes the human heregulin-like factor (HLF) of the  
CC invention. The HLF is involved in the regulation of cell growth.  
CC Detection of different levels of expression of the HLF gene can be used  
CC for the diagnosis of disorders, e.g. in the neural system. In  
CC particular, detection of different levels of HLF gene expression in cells  
CC or body fluid of an individual can be used for diagnosing cancer. The  
CC products can also be used in the treatment of disorders involving  
CC abnormal levels of HLF activity.  
XX  
SQ Sequence 536 BP; 141 A; 131 C; 125 G; 120 T; 19 other;

alignment\_scores:  
Quality: 251.00 Length: 49  
Ratio: 5.340 Gaps: 2  
Percent Similarity: 95.918 Percent Identity: 95.918

alignment\_block:  
US-09-877-665-4 x AAX36424 ..

Align seg 1/1 to: AAX36424 from: 1 to: 536

1 HisphelyProCyArgAspLysAspLeuAlaTyrCysLeuAsnAspG 17  
|||  
99 CACTTCACAACTGCTGAGACAAAGACCTGCGTATTCCTCAAGATG 148  
|||  
17 yelysgluyltyrtingllyalaargcysaspglnphelu 47  
|||  
149 GCGAGTCTTTGTGATTCGAAACCCGACCGGATCCCATTAACACTG 198  
|||  
33 GCTGAGGAGGCTACCAAGAGATCGTGTGATCAATTCTG 241  
|||  
199 GTGCAAGAGGCTACCAAGAGATCGTGTGATCAATTCTG 241  
|||

seq\_name: /SIDSL/9cdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA60998

seq\_documentation\_block:  
ID AAA60998 standard; CDNA: 156 BP.  
XX  
AC AAA60998:  
XX  
DT 09-NOV-2000 (first entry)  
XX  
DE Human NDF EGF-like domain derived peptide encoding CDNA SEQ ID NO.2.  
XX  
KW Human; sensory epithelial cell; growth; stimulant; inner ear; EGF;  
KW epithelial growth factor; NDF; heregulin; monoclonal antibody;  
KW adult rat utricular epithelium; ss.  
XX  
OS Homo sapiens.  
XX  
PN US6080845-A.  
XX  
PD 27-JUN-2000.  
XX  
PF 28-JAN-1999; 99US-0238182.  
XX  
PR 05-AUG-1998; 98US-0129549.  
XX  
PA (AMGE-) AMGEN INC.  
PA Carnahan JF;  
XX  
PI WPI: 2000-451229/39.  
XX  
DR P-PSDB; AAB12602.  
XX  
PT Novel monoclonal antibody against adult rat utricular epithelium useful  
PT in study and research with such organs and tissue  
XX  
PS Disclosure; Column 11-12; 12pp; English.

XX The present invention describes a monoclonal antibody (I) deposited with  
 CC the American Type Culture Collection under accession number HB-12598.  
 CC The antibodies are used in study and research with adult rat utricular  
 CC organs and tissues. The present sequence encodes a peptide derived from  
 CC the epithelial growth factor (EGF) like domains of NDF-alpha and  
 CC NDF-beta (members of the NDF/heregulin protein family, which is used in  
 CC the exemplification of the present invention. The peptide acts as a  
 CC growth stimulant for sensory epithelial cells of the inner ear.

XX Sequence 156 BP; 47 A; 31 C; 42 G; 36 T; 0 other;

alignment\_scores: Quality: 116.50 Length: 48  
 Ratio: 3.149 Gaps: 1

Percent Similarity: 77.083 Percent Identity: 33.333

alignment\_block:  
 US-09-877-665-4 x AA060998 ..

Align seg 1/1 to: AA060998 from: 1 to: 156

1 HispHelyProCysArgAspLysAspLeuAlaTyrCysLeuAsnAspG1 17  
 |||::: ||| ::|||::: ::|||::|||::||| ||  
 4 CATCTGTAAATGTGGGAGAGAGAAACTTCTGTGTGATGAGG 53  
 17 yGlucySpheValIleGluThrLeuThrGlySerHisLysHis...CysA 33  
 |||::|||::|||::|||::|||::|||::|||::|||  
 54 GGAGTGTCTCATGTGTGAAGACCTTCAACCCCTCGATACTGTGCA 103  
 33 rGcylsGluGlyTyrGlnGlyValArgCysAspGlnPheLeu 47  
 ::|||::: |||::: |||::: |||::: |||::: |||::: |||  
 104 AGTGCCAACTCGATGACTGAGAGCAAGATGTCAAACTACGTA 147

seq\_name: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:AA062843

seq\_documentation\_block:  
 ID AA062843 standard; DNA; 192 BP.

XX AA062843;

XX 28-JUL-1994 (first entry)

XX EGFL2.

XX Glial growth factor; GGF; heregulin; mitogenesis;  
 KW Schwann cell; tumour; central nervous system;  
 KW epidermal growth factor; EGF; ss.

XX MO9400140-A.

XX 06-JAN-1994.

XX 29-JUN-1993; 93WO-US06228.

XX 30-JUN-1992; 92US-0907138.

XX 03-SEP-1992; 92US-0940389.

XX 23-OCT-1992; 92US-0965173.

XX 24-MAR-1993; 93US-0036555.

XX (CAMP-) CAMBRIDGE NEUROSCIENCE.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Chen MS, Goodearl A, Hiles I, Marchionni M, Minghetti L;  
 PI Stroobant P, Waterfield M;

XX WPI: 1994-025882/03.  
 XX P-PSDB; AAR55659.

XX Glial mitogenic polypeptide factors - useful for stimulating  
 PT glial cell mitogenesis and treating glial cell tumours  
 XX

PS Claim 53; Fig 39; 178pp; English.

XX EGFL1, EGFL2, EGFL3, EGFL4, EGFL5 and EGFL6 are used  
 CC for the stimulation of glial cell mitogenesis in vivo  
 CC and in vitro.

XX Sequence 192 BP; 61 A; 39 C; 50 G; 42 T; 0 other;

alignment\_scores: Quality: 113.50 Length: 46  
 Ratio: 3.338 Gaps: 1

Percent Similarity: 73.913 Percent Identity: 34.783

alignment\_block:  
 US-09-877-665-4 x AA062843 ..

Align seg 1/1 to: AA062843 from: 1 to: 192

1 HispHelyProCysArgAspLysAspLeuAlaTyrCysLeuAsnAspG1 17  
 |||::: ||| ::|||::: ::|||::|||::||| ||  
 4 CATCTGTCAAGTGTGCAGAGAGAGAAACTTCTGTGTGATGAGG 53  
 17 yGlucySpheValIleGluThrLeuThrGlySerHisLysHis...CysA 33  
 |||::|||::|||::|||::|||::|||::|||::|||  
 54 CGAGTGTCTCATGTGTGAAGACCTTCAAAACCCCTCAAGATACTGTGCA 103  
 33 rGcylsGluGlyTyrGlnGlyValArgCysAspGln 45  
 ::|||::: |||::: |||::: |||::: |||::: |||::: |||  
 104 AGTGCCAACTCGATGACTGAGAGCAAGATGTCAAACTACGTA 141

seq\_name: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:AA058324

seq\_documentation\_block:  
 ID AA058324 standard; DNA; 192 BP.

XX AA058324;

XX 28-JUL-1994 (first entry)

XX EGFL2.

XX Glial growth factor; GGF; heregulin; mitogenesis;  
 KW Schwann cell; tumour; central nervous system; erbB2 receptor;  
 KW antiproliferative; epidermal growth factor; EGF; ss.

XX MO9403644-A.

XX 17-FEB-1994.

XX 10-AUG-1993; 93WO-US07491.

XX 10-AUG-1992; 92US-0927337.

XX 25-SEP-1992; 92US-0951747.

XX 01-DEC-1992; 92US-0984085.

XX 29-JAN-1993; 93US-0011396.

XX (CAMP-) CAMBRIDGE NEUROSCIENCE INC.

XX Gwynne DI, Marchionni M, McBurney RN;

XX WPI: 1994-065731/08.  
 XX P-PSDB; AAR46918.

XX Glial growth factor DNA encoding numerous polypeptide factors  
 PT used for inhibiting cell proliferation - for treating carcinoma  
 PT and nervous disorders  
 XX Disclosure; Fig 40; 178pp; English.

XX The GGF coding segments include regions with EGF-like homology.  
 CC These EGF-like domains can be required for the activation of  
 CC mitogenesis in the binding reaction between GGF ligands confg.

CC such domains and the erbB2 receptor. Pref. antiproliferative  
 CC factors are those which lack these EGF-like domains.  
 XX  
 SO Sequence 192 BP; 61 A; 39 C; 50 G; 42 T; 0 other;

alignment\_scores:  
 Quality: 113.50 Length: 46  
 Ratio: 3.338 Gaps: 1  
 Percent Similarity: 73.913 Percent Identity: 34.783

alignment\_block:  
 US-09-877-665-4 x AAQ58324 ..

Align seg 1/1 to: AAQ58324 from: 1 to: 192.

```

1 H1SPHELYSPROCYSAARGASPLYSAPLEUALATYRCYSLAUNASPG1 17
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
4 CATCTTGCAAGTGTGCAGAGAGAGAAACTTCTGTGTGAATGAGG 53
17 YGLUCSPHEVALLEG1UTRLEUTHRGlySerHisLysHis...CysA 33
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
54 CGAGTGTTCATGCTGGAAGAACCTTCAAAATCCCTCAAGATACCTGTGCA 103
33 TGCYSLYSGLUGLYTYRGInglyValAlaArgCysAspGln 45
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
104 AGTGCCAACTCGATTCACCTGAGAGAGAGATGTACTGAG 141

```

seq\_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAQ74915

seq\_documentation\_block:

ID AAQ74915 standard; DNA: 192 BP.

XX AC AAQ74915;

XX DT 15-AUG-1995 (first entry)

XX DE Human epidermal like growth factor 2 (EGFL2).

XX KW Epidermal like growth factor 2; mammalian muscle cell treatment;

XX KM skeletal; cardiac; smooth; acetylcholine receptor deficiency;

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT mat\_peptide 1..189

XX FT /\*tag- a

XX PN WO9426298-A.

XX PD 24-NOV-1994.

XX PF 06-MAY-1994; 94WO-US05083.

XX PR 06-MAY-1993; 93OS-0059022.

XX PR 08-MAR-1994; 94US-0209204.

XX PA (CAMP-) CAMBRIDGE NEUROSCIENCE.

XX PI Gwynne DI, Marchionni M, Sklar R;

XX DR MPI; 1995-006353/01.

XX DR P-PSDB; AAR67250.

XX PT Treating mammalian muscle diseases and disorders - by admin. of

XX PT GGF2 and other specified polypeptide(s) which bind the p185erbB2

XX PT receptor.

XX PS Clalm 34; Pages 148-149; 241pp; English.

XX CC AAQ74915 encodes AAR67250 human epidermal like growth factor 2 (EGFL2).

XX CC The glial cell mitogenic activity of EGFL2 can be used to treat a

CC variety of mammalian skeletal, cardiac and smooth muscle diseases,  
 CC including acetylcholine receptor deficiency.  
 XX  
 SO Sequence 192 BP; 61 A; 39 C; 50 G; 42 T; 0 other;

alignment\_scores:  
 Quality: 113.50 Length: 46  
 Ratio: 3.338 Gaps: 1  
 Percent Similarity: 73.913 Percent Identity: 34.783

alignment\_block:  
 US-09-877-665-4 x AAQ74915 ..

Align seg 1/1 to: AAQ74915 from: 1 to: 192

```

1 H1SPHELYSPROCYSAARGASPLYSAPLEUALATYRCYSLAUNASPG1 17
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
4 CATCTTGCAAGTGTGCAGAGAGAGAAACTTCTGTGTGAATGAGG 53
17 YGLUCSPHEVALLEG1UTRLEUTHRGlySerHisLysHis...CysA 33
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
54 CGAGTGTTCATGCTGGAAGAACCTTCAAAATCCCTCAAGATACCTGTGCA 103
33 TGCYSLYSGLUGLYTYRGInglyValAlaArgCysAspGln 45
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
104 AGTGCCAACTCGATTCACCTGAGAGAGAGATGTACTGAG 141

```

seq\_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAT48082

seq\_documentation\_block:

ID AAT48082 standard; DNA: 192 BP.

XX AC AAT48082;

XX DT 01-JUN-1997 (first entry)

XX DE EGFL2 DNA.

XX KW Neuregulin; EGFL2; glial growth factor; GGF; heregulin;

XX KM acetylcholine receptor inducing activity; ARIA;

XX KM neu differentiation factor; erbB receptor; mitogen; eye; retina;

XX KM retinopathy; retinitis pigmentosa; macular degeneration;

XX KM wound repair; ss.

XX OS Not specified.

XX PN WO9630403-A1.

XX PD 03-OCT-1996.

XX PF 27-MAR-1996; 96WO-US04240.

XX PR 27-MAR-1995; 95US-0411295.

XX PA (BERN/) BERNINGHAM-MCDONOGH O.

XX PA (CAMP-) CAMBRIDGE NEUROSCIENCE INC.

XX PA (GWYNNE/) GWYNNE D I.

XX PA (MAHA/) MAHANTHAPPA N K.

XX PA (MARC/) MARCHIONNI M A.

XX PA (MCCA/) MCCABE K L.

XX PA (REHT/) REH T A.

XX PI BERNINGHAM-MCDONOGH O, Gwynne DI, Mahanthappa NK;

XX PI Marchionni MA, McCabe KL, Reh TA;

XX DR MPI; 1996-455276/45.

XX DR P-PSDB; AAM09363.

XX PT Use of neuregulin polypeptide(s) - for the treatment of retinal

XX PT cells, partic. for treating retinal diseases or retinal degeneration

XX PS Disclosure; Fig 19; 102pp; English.

XX PS

XX DNA sequences (AAT48081-86) respectively code for EGF1, EGF2, EGF3, EGF4, EGF5 and EGF6 (AAW0362-67) that bind with the extracellular receptors of retinal cells. These polypeptides are useful in new therapeutic approaches to injury or diseases of the eye, esp. degenerative diseases of the retina, based on the promotion of retinal cell function.

Sequence 192 BP; 61 A; 39 C; 50 G; 42 T; 0 other:

alignment\_scores:

Quality:	113.50	Length:	46
Ratio:	3.338	Gaps:	1
Percent Similarity:	73.913	Percent Identity:	34.783

alignment\_block:

US-09-877-665-4 x AAT48082 ..

Align seg 1/1 to: AAT48082 from: 1 to: 192

```

1 HisphelysProCysArgAspLysAspLeuAlaTyrCysLeuAsnAspG1 17
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||
4 CATCTGTCAAGTGTGCAGACAGAGAAACTTCTGTGTGANTGAGG 53
17 yGlucysPheValIleGluThrLeuThrGlySerHisLysHis...CysA 33
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||
54 CGAGTGTCTATGCTGGAAGACCTTCAATCCCTCAAGATCTGTGCA 103
33 rGcysLysGluGlyTyrGlnGlyValArgCysAspGln 45
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||
104 AGTGCCAACTGATTCACGTGAGCGAGATGACTGAG 141

```

seq\_name: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:AAT30990

seq\_documentation\_block:

ID AAT30990 standard; cDNA; 192 BP.

AC AAT30990;

XX 27-AUG-1996 (first entry)

DE Epidermal growth factor-like domain GGF peptide EGF2 coding sequence.

XX G1a1 growth factor; GGF; human; hGF2; Schwann cell; mitogenesis; GGF-1; glial cell; therapy; neuroregulin; herregulin; cellular communication; hGF2; vertebrate; neuron; muscle cell; nervous system; neurotrophic agent; matrix molecule; protease; protease inhibitor; neurological disorder; peripheral neuropathy; amyotrophic lateral sclerosis; nerve injury; spinal muscular atrophy; Alzheimer's disease; spinal cord injury; epidermal growth factor; ss.

OS Synthetic.

XX MO9615812-A1.

PN 30-MAY-1996.

PD 16-NOV-1995; 95MO-US14974.

XX PF 17-NOV-1994; 94US-0341018.

XX PR (CAMEB-) CAMBRIDGE NEUROSCIENCE INC.

XX Bermingham-mcdonogh O, Goldin SM, Gwynne DI, Mahanthappa NK; Marchionni MA, Mcburney RN;

XX WPT. 1996-268341/27.

DR P-PSDB: AAR96076.

XX Use of neuroregulin to affect cellular communication in a vertebrate - by admin. of a neuroregulin or neuroregulin DNA, used to treat neurological disorders, e.g. Alzheimer's disease, peripheral

PT neuropathy etc.

XX Disclosure; Fig 19; 11pp; English.

XX AAT30989-T30994 represent the coding sequences for glial growth factor (GGF) peptides which encompass the epidermal growth factor-like domain (EGFL). These sequences can be used in the method of the invention.

CC administration of neuroregulin or a neuroregulin-producing cell, where the neuroregulin interacts with a first cell type to form product A, which in turn affects a function of a second cell type. The effect on the second cell type results in the production of product B which affects the first (or a third) cell type. Alternatively a neuroregulin coding sequence is incorporated into the genome of a vertebrate cell, and the neuroregulin is then expressed by the cell. The first cell type is a nervous system support cell, preferably a schwann cell, neuron or muscle cell. The second and third cell types are nervous system or muscle cells. The products A and B are neurotrophic agents, matrix molecules, proteases or protease inhibitors. The method can be used to treat neurological disorders, such as peripheral neuropathy, amyotrophic lateral sclerosis, spinal muscular atrophy, nerve injury, Alzheimer's disease or spinal cord injury. It can also induce muscle mitogenesis. This method preferably induces more than one neurotrophic product, and is more effective than using a single neurotrophic factor, which would tend to have a restricted effect on a specific neuronal subtype.

Sequence 192 BP; 61 A; 39 C; 50 G; 42 T; 0 other:

alignment\_scores:

Quality:	113.50	Length:	46
Ratio:	3.338	Gaps:	1
Percent Similarity:	73.913	Percent Identity:	34.783

alignment\_block:

US-09-877-665-4 x AAT30990 ..

Align seg 1/1 to: AAT30990 from: 1 to: 192

```

1 HisphelysProCysArgAspLysAspLeuAlaTyrCysLeuAsnAspG1 17
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||
4 CATCTGTCAAGTGTGCAGACAGAGAAACTTCTGTGTGANTGAGG 53
17 yGlucysPheValIleGluThrLeuThrGlySerHisLysHis...CysA 33
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||
54 CGAGTGTCTATGCTGGAAGACCTTCAATCCCTCAAGATCTGTGCA 103
33 rGcysLysGluGlyTyrGlnGlyValArgCysAspGln 45
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||
104 AGTGCCAACTGATTCACGTGAGCGAGATGACTGAG 141

```

seq\_name: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:AAT06734

seq\_documentation\_block:

ID AAT06734 standard; DNA; 192 BP.

AC AAT06734;

XX 05-JUL-1996 (first entry)

DE Epidermal growth factor-like domain GGF peptide EGF2 coding sequence.

XX G1a1 growth factor; GGF; human; hGF2; Schwann cell; mitogenesis; GGF-1; glial cell; therapy; peripheral nerve damage; demyelination; bovine; neurodegenerative disorder; neural regeneration; acetylcholine receptor; fibroblast proliferation; wound repair; multiple sclerosis; glial tumour; GGF-II; epidermal growth factor; hGF; ss.

OS Synthetic.

XX MO9532724-A1.

PN 07-DEC-1995.



XX MO9426298-A.  
XX 24-NOV-1994.  
XX  
XX 06-MAY-1994; 94MO-US05083.  
XX  
XX 06-MAY-1993; 93US-0059022.  
PR 08-MAR-1994; 94US-0209204.  
XX  
XX  
XX (CAME-) CAMBRIDGE NEUROSCIENCE.  
XX  
XX Gwynne DI, Marchionni M, Sklar R;  
PI  
DR WPI; 1995-006353/01.  
DR P-PSDB; AAR67254.  
XX  
XX  
PT Treating mammalian muscle diseases and disorders - by admin. of  
PT GGF2 and other specified polypeptide(s) which bind the p155erbB2  
PT receptor.  
XX  
XX  
PS Claim 38; Page 151; 241pp; English.  
XX  
XX AAQ7919 encodes AAR67254 human epidermal like growth factor 6 (EGFL6).  
XX The glial cell mitogenic activity of Egfl6 can be used to treat a  
XX variety of mammalian skeletal, cardiac and smooth muscle diseases,  
XX including acetylcholine receptor deficiency.  
XX  
XX Sequence 252 BP; 78 A; 54 C; 63 G; 57 T; 0 other;

```
alignment_scores:
  Quality: 113.50
  Ratio: 3.338
  Percent Similarity: 73.913
  Length: 46
  Gaps: 1
  Percent Identity: 34.783
```

alignment\_block:  
US-09-877-665-4 x AA074915

Align seg 1/1 to: AAQ74919 from: 1 to: 252

[illegible]

**THIS PAGE BLANK (USPTO)**

---

OM of: US-09-877-665-4 to: EST:\* out\_format : pfs  
Date: Sep 15, 2002 12:55 PM

About: Results were produced by the Gencore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

-MODEL=frame-p2n.model -DEV=x1h  
-O=/cgn2.1/USPTO-spool/US09877665/runat.13092002.083912.15645/app-query.fasta.1.103  
-DB=EST -FASTA=fastap -SUPERX=p2n.rst -GAPOP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -OGAPOP=4.500  
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
-GGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000  
-DELEXT=7.000 -START=1 -MATRIX=blomsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500  
-MINLEN=0 -MAXLEN=2000000000 -USER=US09877665.GCEN1.1.2394  
-NCPU=6 -ICPU=3 -LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30  
-NO\_XLPXY -WAIT -THREADS=1

## Search information block:

Query: US-09-877-665-4  
Query length: 47  
Database: EST\*  
Database sequences: 13736207  
Database length: 184157050  
Search time (sec): 1760.310000

## Score list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_est1.AL534571	+	271.00	615.80	4.7e-25	785	AL534571 AL534571 LTI_FL013.FBn1
gb_gss.CNS05BVV	+	167.00	371.93	1.8e-11	1001	AL330340 Tetradon nigroviridis
gb_est2.BM142279	+	113.50	253.45	7.2e-05	570	BM142279 lf33c10.y1 Melton North
gb_est2.BF030411	+	113.50	249.21	0.0001	866	BR304411 601558480F1 NIH.MGC.58
gb_est2.BG150472	+	111.50	247.60	0.0002	642	BR150472 7K01e09.x1 NCI.CGAP.GC
gb_est2.BA238077	+	110.50	247.61	0.0002	510	BF937937 fm70e03.x1 NCI.CGAP.GC
gb_est1.BB637211	+	110.00	243.36	0.0002	568	AA238077 BM10e03.y1 Soares mous
gb_est1.BB637211	+	110.00	243.36	0.0002	652	BB637211 BM637211 RIKEN full-16
gb_est2.B1908144	+	104.50	231.12	0.0013	657	BI908144 6030657448F1 NIH.MGC.11
gb_est2.B1907799	+	104.50	228.14	0.0018	881	BI907799 603065733F1 NIH.MGC.11
gb_hic.BC006492	+	104.50	221.53	0.0043	1692	BC006492 Homo sapiens: myelin
gb_est1.AA968077	+	104.00	237.03	0.0006	327	AA968077 un05h01.r1 Soares mous
gb_est2.BE648742	+	103.50	235.01	0.0008	336	BE648742 UT-M-BH2.2-ac1-d-09-0-
gb_est1.A1137081	+	103.50	233.50	0.0009	413	A1137081 ud5f02.r1 Soares-NMPu
gb_est1.AA798243	+	103.50	231.50	0.0012	503	AA798243 vY05e02.r1 Stratigene
gb_est2.BF941240	+	103.50	231.05	0.0013	526	BF941240 7d95f08.x1 Lupski dors
gb_est2.BE648780	+	101.00	232.43	0.0011	259	BE648780 UT-M-BH2.2-acp-b-12-0-
gb_est1.AV851999	+	97.00	216.97	0.0077	476	AV851999 AV851999 Nori Satoh un
gb_est2.B144093	+	97.00	210.67	0.0173	887	BI464093 603302870F1 NIH.MGC.97
gb_est2.BG133584	+	96.50	211.99	0.0146	694	BG133584 RST13196 Athysys RAGH
gb_est2.BG195532	+	95.50	211.90	0.0148	557	BG195532 RST14725 Athysys RAGH
gb_est1.BB381556	+	95.50	210.78	0.0171	622	BB381556 BB381556 RIKEN full-16
gb_gss.CNS0272C	+	95.50	207.36	0.0265	872	AA213825 Tetradon nigroviridis
gb_est1.BB489119	+	93.50	213.66	0.0119	296	BB489119 BB489119 RIKEN full-16
gb_est1.AV904635	+	93.00	203.93	0.0412	690	AV904635 AV904635 Nori Satoh un
gb_est2.BG196928	+	93.00	201.60	0.0555	868	BI919628 RST16159 Athysys RAGH
gb_est2.BG197944	+	91.50	212.80	0.0132	204	BI919794 RST17195 Athysys RAGH
gb_est1.AV878992	+	91.00	201.22	0.0582	570	AV878992 AV878992 Nori Satoh un
gb_est1.AV880643	+	91.00	200.13	0.0670	635	AV880643 AV880643 Nori Satoh un
gb_est1.AW762061	+	90.00	207.01	0.0277	256	AW762061 ut33c01.y1 NCI.CGAP.ME
gb_est2.B1651936	+	90.00	199.66	0.0712	529	AA476657 ucq79e01.y1 NCI.CGAP.LU
gb_est2.B1410528	+	90.00	193.73	0.1208	795	BI451936 603298677F1 NIH.CGAP.LU
gb_est1.AV897556	+	89.00	193.60	0.1521	949	BI410828 602995334F1 NCI.CGAP.LU
gb_est1.AV897556	+	89.00	193.60	0.1548	765	AV897556 AV897556 Nori Satoh un
gb_est1.AV853430	+	89.00	193.23	0.1622	793	AV853430 AV853430 Nori Satoh un
gb_est2.BG343733	+	88.00	188.28	0.3062	1028	BI343733 BMSME0006016F Horded
gb_est1.A1478167	+	87.50	192.00	0.1880	630	A1478167 tm44h09.x1 NCI.CGAP.KI
gb_est2.B1145755	+	87.00	191.83	0.1942	576	BI145752 602909246F1 NCI.CGAP.LU
gb_est1.AV616555	+	87.00	191.30	0.2079	607	AV616555 AV616555 Bos taurus ov
gb_est1.AA990660	+	87.00	190.48	0.2309	658	AA990660 LD344470.5pTime LD Dros

gb\_est1.BB646928 + 87.00 189.77 0.2530 706 1 BB646928 BB646928 RIKEN full1  
gb\_est2.BG261805 + 87.00 189.61 0.2581 717 1 BG261805 602373601F1 NIH.MGC  
gb\_est2.B1330162 + 87.00 189.17 0.2732 749 1 B1330162 602983780F1 NCI\_CGA  
gb\_est2.BF237257 + 87.00 188.77 0.2875 779 1 BF237257 602025418F1 NCI\_CGA  
gb\_est2.B1148137 + 87.00 188.20 0.3093 824 1 B1148137 602912309F1 NCI\_CGA

seq\_name: gb\_est1:AL534571

seq\_documentation\_block:

LOCUS AL534571 785 bp mRNA linear EST 13-FEB-2001  
DEFINITION AL534571 LTI\_FL013.FBn1 Homo sapiens CDNA clone CS0DF0041618.5  
ACCESSION AL534571  
VERSION AL534571.1 GI:12798064  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

## FEATURES

source  
1..785  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DF0041618"  
/dev\_stage="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)"  
/lab\_host="DH10B"  
/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dt) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville , Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

## BASE COUNT

222 a 161 c 181 g 218 t 3 others

## ALIGNMENT

alignment\_scores:  
Quality: 271.00 Length: 47  
Ratio: 5.891 Gaps: 0  
Percent Similarity: 97.872 Percent Identity: 97.872  
alignment\_block:  
US-09-877-665-4 x AL534571 ..  
Align seg 1/1 to: AL534571 from: 1 to: 785

1 HisPhelysProCyAtGaSpLyAsPleuLaTyCySleuSaAspG1 17  
|||||  
459 CACTTCAAACCCCTCCGACAAAGGACCTTCATCTCATATGATGG 508  
|||||

17 YGluCyBheValIleGluBhrLeuPhrGlySerHisLysHisCysArgC 34  
|||||  
509 CGAGTGTCTTGTATCGAMACCTGACGGATCCCATTAACACTGTCTGGT 558  
|||||

34 YslYsgLugLyTyGInGlyValArgCysAspGlnPheLeu 47  
|||||  
559 GCAAGAAGAGCTACCAAGAGTGCTGTGTATCAATTTCTG 599  
|||||

seq\_name: gb\_gss:CNS05BVV



```

seq_documentation_block:
LOCUS      CNS05BVY              1001 bp    DNA             linear    GSS 26-MAY-2000
DEFINITION Tetracodon nigroviridis genome survey sequence T3 end of clone
            012D17 of library C from Tetracodon nigroviridis, genomic survey
            sequence.
ACCESSION  AL330340.1 GI:8223962
VERSION    AL330340
KEYWORDS   GSS: genome survey sequence.
SOURCE     Tetracodon nigroviridis.
            Tetracodon nigroviridis.
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetracodontiformes;
            Tetracodontidae; Tetracodon.
REFERENCE  1 (bases 1 to 1001)
AUTHORS   Roest-Crollius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
TITLE      Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetracodon nigroviridis
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 1001)
AUTHORS   Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
TITLE      Human gene number estimate provided by genome wide analysis using
            Tetracodon nigroviridis DNA sequence
JOURNAL    Unpublished
REFERENCE  3 (bases 1 to 1001)
AUTHORS    Genoscope.
TITLE      Direct Submission
COMMENT     Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
            This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetracodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetracodon.
FEATURES   source
            1..1001
               /organism="Tetracodon nigroviridis"
               /db_xref="taxon:99883"
               /clone_1lb="012D17"
               /clone_1lb="C"
               /note="Genoscope sequence ID : C0AC012CB09T1-end : T3"
BASE COUNT      272 a      218 c      219 g      278 t      14 others
ORIGIN
alignment_scores:
            Quality: 167.00      Length: 33
            Ratio: 5.387      Gaps: 0
            Percent Similarity: 93.939      Percent Identity: 81.818
alignment_block:
US-09-877-665-4 x CNS05BVY
Align seg 1/1 to: CNS05BVY from: 1 to: 1001
1 HisphelysProcysargaspLysaspLeuAlaTyrCysLeuAsnAspG1 17
|||||
667 CATTTCAGCCCTGCATGAGACGACCTGCGTACTGTGAGACGCTGG 716
|||||
17 yglucyPpheValIleGIuThrLeuThrGlySerHisLysHisGysArg 33
|||||
717 AGAGTGTCTTCATCGAGACGCTCAGCGGCGCTCACAAACACTGCAAG 765
|||||
seq_name: gb_est2:BM142279
seq_documentation_block:
LOCUS      BM142279              570 bp    mRNA             linear    EST 29-NOV-2001
DEFINITION If33c10.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
            cDNA 5' similar to TR:035947 035947 PRO-NEUREGULIN-1 PRECURSOR ;,
            mRNA sequence.

```

```

ACCESSION  BM142279              GI:17152346
VERSION    BM142279.1
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 570)
AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scaer,C., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Maira,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritzer,E., Ronko,I., Bennett,D., Cardenas
            M., Gibbons,M., McCann,R., Cole,R., Tsagaratshvili,R., Williams,T.,
            Jackson,Y. and Bowers,Y.
TITLE      Endocrine Pancreas Consortium
JOURNAL    Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Douglas Melton DNA sequencing by:
            Washington University Genome Sequencing Center for information on
            obtaining a clone please contact: Juliana Brown
            (brownjefas.harvard.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 429.
FEATURES   source
            1..570
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone_1lb="Melton Normalized Human Islet 4 N4-HIS 1"
               /sex="Both"
               /tissue.type="Islets of Langerhans"
               /dev_stage="Adult"
               /lab_host="DH10B"
               /note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
               Site_2: Sal 1; Starting library constructed using
               Superscript plasmid library kit (Life Technologies). cDNA
               made by oligo-dT priming. Size-selected by column
               fractionation; average insert size 1.08 kb. Library was
               amplified once on solid support and plasmid DNA from
               library was prepared. The library DNA was normalized by
               method #4 from Bonaldo, Lennon, and Soares 1996 Genome
               Research 6:791-806; 0.5 microgram single-stranded library
               plasmid DNA was mixed with 5 micrograms PCR product
               representing library inserts and hybridized to an EcoT of
               20. Single-stranded (unhybridized) plasmids were isolated
               by hydroxyapatite chromatography and used to make this
               library."
BASE COUNT      175 a      140 c      132 g      123 t
ORIGIN
alignment_scores:
            Quality: 113.50      Length: 46
            Ratio: 3.338      Gaps: 1
            Percent Similarity: 73.913      Percent Identity: 34.783
alignment_block:
US-09-877-665-4 x BM142279
Align seg 1/1 to: BM142279 from: 1 to: 570
1 HisphelysProcysargaspLysaspLeuAlaTyrCysLeuAsnAspG1 17
|||||
157 CATTTCATAAATGTGCGAGAGAGAAAACCTTGTGTGTGGAATGAGAG 206
|||||
17 yglucyPpheValIleGIuThrLeuThrGlySerHisLysHis...CysA 33
|||||

```





Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.



s. 6 (9), 791-806 (1996)

JOURNAL  
Genome Res. 6 (9), 791-806 (1996)  
MEDLINE  
97044477

Contact: ChIn, H

National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 0000

Email: [mestem@nlh.gov](mailto:mestem@nlh.gov)  
 Fax: 302-843-5650  
 cDNA library Preparation: M.B. Soares Lab Clone distribution:  
 Researchers may obtain BMP cDNA clones from RESEARCH GENETICS. It  
 should be noted that Bento Soares is generating a small number of  
 additional specialized non-redundant arrays of BMP cDNAs whose  
 availability will be considered under appropriate and limited  
 collaborative arrangements  
 seq primer: M13 Reverse.

### Source

BASE COUNT  
ORIGIN  
97 a 106 c 74 g 79 t

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_uid="M-BH2.2-aol-d-09-0-01"
/clone_lib="NIH_BMAP_M_S3.2"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/notice="Vector: pUT730-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; The
NIH_BMAP_M_S3.2 library is a subtracted library of a
series, ultimately derived from a mixture of individually
tagged normalized libraries from ten regions of the mouse
brain (cerebellum, brain stems, olfactory bulbs,
hypothalamus, cortex, amygdala, basal ganglia, pineal
gland, striatum, hippocampus) after a series of
subtractions to reduce the representation of cDNAs from
which ESTs had already been generated. The following
serially subtracted libraries were generated in this
process: NIH_BMAP_M_S3.2, NIH_BMAP_M_S2, NIH_BMAP_M_S1.
The subtracted library (NIH_BMAP_M_S3.2) was constructed
as follows: PCR amplified cDNA inserts from NIH_BMAP_M_S2
clones from which 3' ESTs had been derived was used as a
driver in a hybridization with the NIH_BMAP_M_S2 library
in the form of single-stranded circles. The remaining
single-stranded circles (subtracted library) was purified
by hydroxyapatite column chromatography, converted to
double-stranded circles and electroporated into DH10B
bacteria (Life Technologies) to generate the
NIH_BMAP_M_S3.2 library. This procedure has been
previously described (Donaldso, Lennon and Soares, Genome
Research 6:791-806, 1996)."

```

US-09-877-665-4 X BE648742

Quality:	103.50	Length:	48
Ratio:	2.957	Gaps:	1
Percent Similarity:	72.917	Percent Identity:	31.250

Align seg 1/1 to: BE648742 from: 1 to: 356

1 Hisphelybriocystatrypsaspheulayrcysleuanaasppl 17  
187 CACCTCATTAAGGTGCGGAGAGAGAGAAACTTCTGTGTGAAAGG 230  
17 ygiucyspheallieglthtrheuglyserlisyshs...cysa 33  
|||||  
237 CGAGTCTCTTAAGGTGAGAGACACTCTCAAAACCCCTCAAGATCTTGCA 286  
33 rgcysylslnluyrllyrlnglyvalargcysasbplnphelu 47  
ACGTCCTCAATATAGTTACTGTGTGATCTGTCCCAAAACTACGTA 330

seq\_name: gb\_est1:AI197081

seq\_documentation\_block

LOCUS	AI197081	413 bp	mRNA	linear	EST 14-OCT-1998
DEFINITION	ud57c02.r1 Soares_NMPu Mus musculus cdna clone IMAGE:1449986 5'				

similar to TR:Q07111 Q07111 GLIAL GROWTH FACTOR 2 PRECURSOR ;, mRNA sequence.

ACCESSION	AI197081
VERSION	AI197081.1
GI	GI:3749687

**SOURCE** house mouse.

**ORGANISM** *Mus musculus*

REFERENCE  
1 (bases 1 to 413)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus

## AUTHORS

Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, D.

TITLE	JOURNAL
The WashU-HIMI Mouse EST Project	Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800  
Fax: 314 286 1810

Email: [mouseeest@watson.wustl.edu](mailto:mouseeest@watson.wustl.edu)

This clone is available royalty-free through LNL; contact the IMAGE Consortium ([info@image.lnl.gov](mailto:info@image.lnl.gov)) for further information.  
MGI:923302

Seq primer: -28ml3 rev2 ET from Amersham  
High quality sequence stop: 412.

## FEATURES

## Source

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1449986"
/clone_lib="Soares_NMPu"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pRT73D-Pac (Pharmacia) with
a modified polylinker; 1st strand cDNA was prepared from
pregnant mouse uterus, and was then primed with a Not I -
Oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pRT73
vector. Library is normalized. Library was constructed by
Bento Soares and M Fatima Bonaldo."

```

ORIGIN

109 9 91 3 507

alignment\_scores:

	70T	R TS	COT	H	N
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					
36					
37					
38					
39					
40					
41					
42					
43					
44					
45					
46					
47					
48					
49					
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					
61					
62					
63					
64					
65					
66					
67					
68					
69					
70					
71					
72					
73					
74					
75					
76					
77					
78					
79					
80					
81					
82					
83					
84					
85					
86					
87					
88					
89					
90					
91					
92					
93					
94					
95					
96					
97					
98					
99					
100					

alignment\_block:

US-09-877-665-4 x A1197081

Align seg 1/1 to: A1197081 from: 1 to: 413

17 ygiuCySPhevalIIegluThrLeuThyGlySerHisLysHis...Cysa 33  
 218 CGAGTCTTCACAGGAGGAAAGACACTGTCAAAACCCCTCAAGATACCTTGCA 267  
 168 CACTCATTAAGAGTGTCCGAGAGGAAGAAACTTCTGTGTGAATGACG 212  
 1 hasPheLysPheIrcCysATGAspLysAspLeuAlaATGTCysLeuAlaAspGI 17  
 1 hasPheLysPheIrcCysATGAspLysAspLeuAlaATGTCysLeuAlaAspGI 17





**THIS PAGE BLANK (USPTO)**

---

OM of: US-09-877-665-4 to: GenEmbl: \* out\_format : pfs  
Date: Sep 15, 2002 1:28 PM  
About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:  
-MODEL-frame.p2n.model -DEV-x1h  
-O/-cgn2.1/USPTO.spool/US09877665/Runat\_13092002.083912.15659/app-query.fasta.1.103  
-DB-GenEmbl -OPMT-fastap -SUFFIX-p2n.rge -GAPOP-12.000  
-GAPEXT-4.000 -MINMATCH-0.100 -LOOPEXT-0.000  
-GAPOP-4.500 -GAPEXT-0.050 -XGAPOP-10.000 -XGAPEXT-0.500  
-XGAPOP-6.000 -DELEXT-7.000 -XGAPOP-10.000 -XGAPEXT-0.500  
-DELOP-6.000 -DELEXT-7.000 -START-1 -MATRIX-biosum62  
-TRANS-numa40.cdi -LIST-45 -DOCALLIGN-200 -THR\_SCORE-pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN-15 -MODE-LOCAL -OUTFMT-pfs  
-NOR-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09877665.cgn1.1.3708 -NCP=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPHY -WAIT -THREADS=1

Search information block:  
Query: US-09-877-665-4  
Query length: 47  
Database: GenEmbl: \*  
Database length: 197656  
Search time (sec): 1997.040000

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
gb_ro:AF010130	277.00	669.10	6.6e-29	2538	AF010130 Mus musculus neuregulin
gb_hng:AC096527	203.00	450.32	1.0e-16	137600	AC096527 Rattus norvegicus c
gb_pr:AL391478	195.00	427.98	1.8e-15	180625	AL391478 Human DNA sequence
gb_ov:AF076618	126.50	289.24	9.5e-08	3105	AF076618 Xenopus laevis neureg
gb_pat:AR02501	116.50	288.69	1.0e-07	156	AR02501 Sequence 2 from patent
gb_pat:AR068979	113.50	279.44	3.3e-07	192	AR068979 Sequence 155 from pat
gb_pat:AR103359	113.50	279.44	3.3e-07	192	AR103359 Sequence 155 from pat
gb_pat:AR133506	113.50	279.44	3.3e-07	192	AR133506 Sequence 155 from pat
gb_pat:AR141835	113.50	279.44	3.3e-07	192	AR141835 Sequence 155 from pat
gb_pat:AR143352	113.50	279.44	3.3e-07	192	AR143352 Sequence 155 from pat
gb_pat:AR151961	113.50	279.44	3.3e-07	192	AR151961 Sequence 155 from pat
gb_pat:136441	113.50	279.44	3.3e-07	192	136441 Sequence 155 from patent
gb_pat:188141	113.50	279.44	3.3e-07	192	188141 Sequence 155 from patent
gb_pat:AR022505	113.50	277.21	4.5e-07	252	AR022505 Sequence 159 from pat
gb_pat:AR068983	113.50	277.21	4.5e-07	252	AR068983 Sequence 159 from pat
gb_pat:AR103363	113.50	277.21	4.5e-07	252	AR103363 Sequence 69 from paten
gb_pat:AR133010	113.50	277.21	4.5e-07	252	AR133010 Sequence 159 from pat
gb_pat:AR141839	113.50	277.21	4.5e-07	252	AR141839 Sequence 159 from pat
gb_pat:AR143356	113.50	277.21	4.5e-07	252	AR143356 Sequence 159 from pat
gb_pat:AR151965	113.50	277.21	4.5e-07	252	AR151965 Sequence 159 from pat
gb_pat:136445	113.50	277.21	4.5e-07	252	136445 Sequence 159 from patent
gb_pat:188145	113.50	277.21	4.5e-07	252	188145 Sequence 159 from patent
gb_pat:AR022504	113.50	276.73	4.7e-07	267	AR022504 Sequence 158 from pat
gb_pat:AR068982	113.50	276.73	4.7e-07	267	AR068982 Sequence 158 from pat
gb_pat:AR103362	113.50	276.73	4.7e-07	267	AR103362 Sequence 67 from paten
gb_pat:AR135009	113.50	276.73	4.7e-07	267	AR135009 Sequence 158 from pat
gb_pat:AR141838	113.50	276.73	4.7e-07	267	AR141838 Sequence 158 from pat
gb_pat:AR143355	113.50	276.73	4.7e-07	267	AR143355 Sequence 158 from pat
gb_pat:AR151964	113.50	276.73	4.7e-07	267	AR151964 Sequence 158 from pat
gb_pat:136444	113.50	276.73	4.7e-07	267	136444 Sequence 158 from patent
gb_pat:188144	113.50	276.73	4.7e-07	267	188144 Sequence 158 from patent
gb_pr:HS02327	113.50	266.87	1.7e-06	886	HS02327 Human clone nbf19 neu d
gb_pat:AR022498	113.50	264.80	2.2e-06	1140	AR022498 Sequence 149 from pat
gb_pat:AR068976	113.50	264.80	2.2e-06	1140	AR068976 Sequence 149 from pat
gb_pat:AR103355	113.50	264.80	2.2e-06	1140	AR103355 Sequence 55 from pat
gb_pat:AR135003	113.50	264.80	2.2e-06	1140	AR135003 Sequence 149 from pat
gb_pat:AR141832	113.50	264.80	2.2e-06	1140	AR141832 Sequence 149 from pat
gb_pat:AR151958	113.50	264.80	2.2e-06	1140	AR151958 Sequence 149 from pat

gb_pat:136438	+	113.50	264.80	2.2e-06	1140	136438 Sequence 149 from pa
gb_pat:188138	+	113.50	264.80	2.2e-06	1140	188138 Sequence 149 from pa
gb_pat:A29423	+	113.50	264.43	2.3e-06	1193	A29423 putative bovine GGF-
gb_pat:AR022483	+	113.50	264.43	2.3e-06	1193	AR022483 Sequence 134 from
gb_pat:AR068961	+	113.50	264.43	2.3e-06	1193	AR068961 Sequence 134 from

seq\_name: gb\_ro:AF010130  
seq\_documentation\_block:  
LOCUS AF010130 2538 bp mRNA linear ROD 23-SEP-1997  
DEFINITION Mus musculus neuregulin-3 (NRG3) mRNA, complete cds.  
ACCESSION AF010130  
VERSION AF010130.1 GI:2429163  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 2538)  
Zhang,D., Sliwkowski,M.X., Mark,M., Frantz,G., Akita,R., Sun,Y.,  
Hillan,K., Crowley,C., Brush,J. and Godowski,P.J.  
Neuregulin-3 (NRG3): A novel neural tissue-enriched protein that  
binds and activates ErbB4  
Proc. Natl. Acad. Sci. U.S.A. 94 (18), 9562-9567 (1997)  
JOURNAL 97420720  
MEDLINE 2 (bases 1 to 2538)  
REFERENCE Zhang,D., Sliwkowski,M.X., Mark,M.R., Frantz,G., Akita,R., Sun,Y.,  
Hillan,K., Crowley,C., Brush,J. and Godowski,P.J.  
A Novel Neural Tissue-Enriched Protein That Binds and Activates ErbB4  
Proc. Natl. Acad. Sci. U.S.A. 94 (18), 9562-9567 (1997)  
TITLE Direct Submission  
JOURNAL Submitted (23-JUN-1997) Molecular Biology, Genentech, 1 DNA Way, S.  
San Francisco, CA 94080, USA  
FEATURES  
source location/Qualifiers  
1..2538  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
1..2538  
/gene="NRG3"  
289..2430  
/gene="NRG3"  
/function="ErbB4 ligand"  
/codon\_start=1  
/product="neuregulin-3"  
/protein\_id="AA070914.1"  
/db\_xref="GI:2429164"  
/translation="MSEGAAGASPPGASAAASAEETGTAATAAAAAAGGPGGEG  
AAEPRELRSDICIVNROQTWLCVPEIFIGIGLSMLKTVGVSKEYVPTDL  
VDSKMGDDPEFLSKSSPPEMTTSTTTSTTSTTSTTSTTSTTSTTSTTSTTSTT  
ITRATPRGPHAVPIRASPRSTTAAATTAAPPVLTTPAPSSSTTPGSRPPGAPST  
QAMPSPRAIVATSYSLDSTPSWTLSPPODAASSSPSSSTSTTPETSTSPK  
HTTYSSTERSEHFPCRDLDLAVCLNDECEVETLTGSHKRCRKEGVGRCOPL  
PTDILSDPDHDLGIEFMESDYQROVLSISCIIFGIYVGMCAEFKSKROAK  
QIOELKESQNKSNYSLKASSTKSLKSHVLYONKADHPPTALEKIMESFSA  
POSPEVSPDRSGSQPIKHSFGSGRGLHRTFRAPRSPRGIVGAYOOLE  
SRIPDDIPROGIEVREKTSILTPOLKCVREPLDIKYVNGTIRQNAASINMOPSR  
ETNPTFNLDQKDLVIGTISPRANSVPIITPMQLEETCKOMQFISDKIKKNSYSA  
DIVNASMVSQCLLEQDEVKILLETVOEQRIILTLDARSSEDFELASMETEDSASNT  
AFPLSPAKEREREAQFVLRNEIDQDSVLTK"

CDS	/gene="NR3"	289.2430	/gene="NR3"
	/function="ErB4_ligand"		
	/codon_start=1		
	/product="neuregulin-3"		
	/protein_id="AB70914.1"		
	/db_xref="GI:2429164"		
	/translation="MSGGAAAGSPGASAAASAEAGTAAAAAAAAAAAGGPGGAGC AAPPPELRCSDCIVNNROOTWLCVPLFGFGIGLISMLMLVIYVGSKERVPTDL VBSKGGOPEFLSKSPSPKAMETTTTSTTSPATPSGGASASRTPIRISTRLTT ITAAPRPFGHVRPIRASPRSTAPRANTAPPYLSTTAPFSSSTGSRPMNGASST OAMPSTPAAYATSSYLHSDTPMTLSPPODAAAASSSPSSSTTTPESTSPKE HTTYSSTERSSEHKPCRDMDIACYLNDGEFVETLFGSKHKRCGEGVGVCDDPL PKTDSLSDPTDHGIEFMESDEVYORVSLSTCTFGIYIVGMCAFTFKSKQAK QIOEHLKESONGKNYSLKASSTKSESIMKSHVLOWSKADRIHPVALEIMESSFA POSPEFTVTPDCRSOPIKHHSPOGRSGMLRNTFRAPSPSRSLGGIVPAQOPEE SRIPDDDTIPDCGIEVYKTIISHLPIDLMCEVRLDLKYSVNGSLTOONASINQOLSR ENPNYNSLSDCLDGLVGLSPRANSVPILIPSMGLEPFCMOMGTSIDYKSIKWCNSYA DIYMAFPYSDCLLEBOEQEYKILLETVOEDRIILTLPARSEDEFLASMETEDASENT AFPLSPSTKSEAEQFVLNKEQROSVLTK"		
BASE COUNT	606 a	785 c	649 g
ORIGIN			498 t
alignment_scores:	Quality: 277.00	Length: 47	
	Ratio: 5.894	Gaps: 0	
Percent Similarity: 100.000	Percent Identity: 100.000		

alignment\_block:  
US-09-877-665-4 x AF010130  
Align seg 1/1 to: AF010130 from: 1 to: 2538  
1 Hisphenylstrocysargaspylsasplealatyrcysleuansapgl 17  
|||||

1150 CACTCAACCTGTCGACACAGACCTGGCGTATTGTCATCATGATGG 1199

17 yglucysphevalilleglthleuthrclyserhislyshiscysargc 34  
 |||||

1200 TGAATGCTTGTGATGAGACCTGACAGATCCATAGACCTGCGGT 1249  
 |||||

34 yslsgsluglytytrglnqlyvalargcysaspglnphelau 47  
 |||||

1250 GCAAGGAAGGCTACCAAGAGTCCGTGTGATCATTTCTG 1290  
 |||||

seq\_name: gb\_htg:AC096527

seq\_documentation\_block:  
 LOCUS AC096527 137600 bp DNA linear HTG 20-DEC-2001  
 DEFINITION Rattus norvegicus clone CH230-78H3, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
 67 unordered pieces.  
 AC096527  
 AC096527 GI:17962756  
 VERSION AC096527.2  
 KEYWORDS HTG; HTGS\_PHASE1.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus  
 1 (bases 1 to 137600)  
 Muzny,D.M., Adams,C., Adio,Oduola,B., Al-Osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Banks,T., Barbara,J.,  
 Benton,J., Bimaga,K., Blankenburg,K., Bohn,D., Bouck,J.,  
 Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C.,  
 Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,  
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,  
 Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,  
 Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,  
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,  
 Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Drepper,H.,  
 Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C.,  
 Elina,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,  
 Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,  
 Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,  
 Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A.,  
 Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,  
 Hollins,B., Homs,F., Howard,S., Huber,J., Huliy,S., Hume,J.,  
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jollivet,S.,  
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,  
 Kovach,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,  
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,  
 Lounsged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,  
 Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,  
 Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,  
 Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,  
 Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,  
 Oguh,M., Okunolu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
 Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,  
 Quiles,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,  
 Slisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,  
 Stone,H., Sutton,A., Sytek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,  
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Tellrod,B., Thomas,N.,  
 Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,  
 Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,  
 Watlington,S., Williams,G., Williamson,A., Wleczky,R., Woodson,S.,  
 Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Nelson,D.,  
 Weinstein,G. and Gibbs,R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 137600)  
 Worley,K.C.  
 Direct Submission  
 Submitted (18-SEP-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Dec 20, 2001 this sequence version replaced gl.15638976.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: GGC1  
 Center clone name: CH230-78H3  
 ----- Summary Statistics  
 Assembly program: Phrap; version 0.990329first call to  
 findPhrapList  
 Consensus quality: 105498 bases at least Q40  
 Consensus quality: 121268 bases at least Q30  
 Consensus quality: 130770 bases at least Q20  
 Estimated insert size: 102221; sum-of-coverage  
 Quality coverage: 0x in Q20 bases; agarose-tp estimation  
 Quality coverage: 1.5x in Q20 bases; sum-of-coverage  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 67 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 1  
 8776: contig of 8775 bp in length  
 8776: gap of unknown length  
 8876: contig of 6705 bp in length  
 15581: gap of unknown length  
 15581: gap of unknown length  
 15581: contig of 3823 bp in length  
 15581: gap of unknown length  
 15581: contig of 3954 bp in length  
 15581: gap of unknown length  
 15581: contig of 3693 bp in length  
 15581: gap of unknown length  
 15581: contig of 2223 bp in length  
 15581: gap of unknown length  
 15581: contig of 2507 bp in length  
 15581: gap of unknown length  
 15581: contig of 3696 bp in length  
 15581: gap of unknown length  
 15581: contig of 4109 bp in length  
 15581: gap of unknown length  
 15581: contig of 2995 bp in length  
 15581: gap of unknown length  
 15581: contig of 2680 bp in length  
 15581: gap of unknown length  
 15581: contig of 2317 bp in length  
 15581: gap of unknown length  
 15581: contig of 1821 bp in length  
 15581: gap of unknown length  
 15581: contig of 1342 bp in length  
 15581: gap of unknown length  
 15581: contig of 2237 bp in length  
 15581: gap of unknown length  
 15581: contig of 2293 bp in length  
 15581: gap of unknown length  
 15581: contig of 2421 bp in length  
 15581: gap of unknown length  
 15581: contig of 1859 bp in length  
 15581: gap of unknown length  
 15581: contig of 2206 bp in length  
 15581: gap of unknown length  
 15581: contig of 2000 bp in length  
 15581: gap of unknown length  
 15581: contig of 1681 bp in length  
 15581: gap of unknown length  
 15581: contig of 2024 bp in length  
 15581: gap of unknown length  
 15581: contig of 1417 bp in length

```

* 71009 71108: gap of unknown length
* 71109 72408: contig of 1300 bp in length
* 72409 72508: gap of unknown length
* 72509 73781: contig of 1273 bp in length
* 73782 73881: gap of unknown length
* 73882 75196: contig of 1315 bp in length
* 75197 75297: gap of unknown length
* 75297 77054: contig of 1758 bp in length
* 77055 77154: gap of unknown length
* 77155 78785: contig of 1631 bp in length
* 78786 78885: gap of unknown length
* 78886 80624: contig of 1738 bp in length
* 80624 80724: gap of unknown length
* 80724 82279: contig of 1555 bp in length
* 82279 82378: gap of unknown length
* 82378 84106: contig of 1728 bp in length
* 84107 84206: gap of unknown length
* 84207 85530: contig of 1324 bp in length
* 85531 85630: gap of unknown length
* 85631 87347: contig of 1717 bp in length
* 87348 87447: gap of unknown length
* 87448 88988: contig of 1541 bp in length
* 88989 90225: gap of unknown length
* 90226 90325: contig of 1137 bp in length
* 90326 91776: gap of unknown length
* 91777 91876: contig of 1451 bp in length
* 91877 93278: gap of unknown length
* 93278 93379: contig of 1402 bp in length
* 93379 94793: gap of unknown length
* 94794 94893: contig of 1415 bp in length
* 94894 96355: gap of unknown length
* 96356 96455: contig of 1462 bp in length
* 96456 97928: gap of unknown length
* 97929 98029: contig of 1473 bp in length
* 98029 99454: gap of unknown length
* 99455 99555: contig of 1426 bp in length
* 99555 101095: gap of unknown length
* 101096 101935: contig of 1541 bp in length
* 10196 102934: gap of unknown length
* 102935 103034: contig of 1739 bp in length
* 103035 104472: gap of unknown length
* 104473 104572: contig of 1438 bp in length
* 104573 106427: gap of unknown length
* 106428 106527: contig of 1855 bp in length
* 106528 108182: gap of unknown length
* 108183 108282: contig of 1655 bp in length
* 108283 110058: gap of unknown length
* 110059 110158: contig of 1776 bp in length
* 110159 111820: gap of unknown length
* 111821 111920: contig of 1662 bp in length
* 111921 113107: gap of unknown length
* 113108 113207: contig of 1187 bp in length
* 113208 114526: gap of unknown length
* 114527 114627: contig of 1319 bp in length
* 114628 116291: gap of unknown length
* 116292 116391: contig of 1665 bp in length
* 116392 117584: gap of unknown length
* 117585 117684: contig of 1193 bp in length
* 117686 118961: gap of unknown length
* 118962 119061: contig of 1277 bp in length
* 119062 120491: gap of unknown length
* 120491 120590: contig of 1429 bp in length
* 120591 121634: gap of unknown length
* 121635 121734: contig of 1044 bp in length
* 121735 123040: gap of unknown length
* 123041 123140: contig of 1306 bp in length

```

alignment\_scores:  
 Quality: 203.00 Length: 46  
 Ratio: 4.951 Gaps: 0  
 Percent Similarity: 89.130 Percent Identity: 76.087

alignment\_block:  
 US-09-877-665-4 x AC096527/rev ..  
 Align seg 1/1 to reverse of: AC096527 from: 1 to: 137600

1 HisphelysProcysArgAspIysAspLeuAlaIATryCysLeuAsnAspG1 17  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||  
 48942 CACTTCAACCTGTCGAGACAAGACCTGGCGTATGTCCTCAAGACGCG 48893

17 yGluCysPheValIleGluThrLeuThrGlySerHisLysHisCysArgC 34  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||  
 48892 TGAATGCTTGTGATGAGACCTTACAGATCCATCCATAGACCTGTGGT 48843

34 ySLysGluIcylGlnGlyAlaIArgCysAspGlnPhe 46  
 ::::::::::::::::::::|||  
 48842 AAGCCCTGACAGCAATAGTTAAAGTGTAGTCTCTT 48805

seg\_name: gb\_pr:AL391478

seg\_documentation\_block:

LOCUS AL391478 180625 bp DNA linear PRI 18-APR-2001

DEFINITION Human DNA sequence from clone RP11-552113 on chromosome 10,  
 complete sequence.

ACCESSION AL391478

VERSION AL391478.14 GI:13750895

KEYWORDS HTG.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 180625)

AUTHORS Clark,S.

TITLE Direct Submmission

JOURNAL Submitted (18-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

COMMENT requests: clonerequests@sanger.ac.uk  
 On Apr 21, 2001 this sequence version replaced gi:13396637.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence substation  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >=  
 30). an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em: EMBL, Sw: SWISSPROT, Tr: TrEMBL, Wp: WORMPEP; Information on the WORMPEP  
 database can be found at  
 http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence  
 was generated from part of bacterial clone contigs of human  
 chromosome 10, constructed by the Sanger Centre Chromosome 10  
 Mapping Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr10  
 RP11-552113 is from the library RPCI-11.2 constructed by the group  
 of Pieter de Jong. For further details see  
 http://www.chori.org/Bacpac/home.htm

VECTOR: pBACe3.6  
 This sequence is the entire insert of clone RP11-552113 The true  
 left end of clone RP11-52203 is at 86094 in this sequence. The true  
 right end of clone RP11-319L5 is at 48826 in this sequence.

#### FEATURES

source

1. 180625  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="10"  
 /clone="RP11-552113"  
 /clone\_11b="RPCI-11.2"

```
repeat_region 355. .665
/note="AluX repeat: matches 1. .310 of consensus"
951. .1018
/note="34 copies 2 mer tt 70% conserved"
repeat_region 1096. .1666
/note="LIM4 repeat: matches 3695. .4279 of consensus"
1793. .2104
/note="AluY repeat: matches 1. .311 of consensus"
3064. .3099
/note="MADE1 repeat: matches 1. .36 of consensus"
3533. .3759
/note="AluSx repeat: matches 1. .283 of consensus"
5297. .5523
/note="L2 repeat: matches 1912. .2153 of consensus"
6335. .6577
/note="3 copies 81 mer 93% conserved"
7016. .7223
/note="MIR repeat: matches 56. .257 of consensus"
8275. .8558
/note="AluSx repeat: matches 10. .297 of consensus"
8559. .8735
/note="AluSg/x repeat: matches 131. .310 of consensus"
9340. .9511
/note="MIR repeat: matches 36. .219 of consensus"
9898. .10207
/note="AluY repeat: matches 1. .308 of consensus"
11032. .11317
/note="AluSx repeat: matches 1. .305 of consensus"
12667. .12793
/note="L2 repeat: matches 2612. .2750 of consensus"
13090. .13343
/note="MER4C repeat: matches 157. .1008 of consensus"
13363. .13721
/note="MER4A repeat: matches 1. .342 of consensus"
14225. .14260
/note="L2 repeat: matches 2714. .2749 of consensus"
16071. .16274
/note="LTR1 repeat: matches 1. .203 of consensus"
16289. .16420
/note="LTR28 repeat: matches 432. .558 of consensus"
16424. .16597
/note="LTR1 repeat: matches 534. .696 of consensus"
18812. .19232
/note="LIM4 repeat: matches 3410. .3826 of consensus"
19233. .19535
/note="AluSg repeat: matches 1. .303 of consensus"
19536. .20534
/note="LIM4 repeat: matches 2393. .3410 of consensus"
20542. .20777
/note="AluY repeat: matches 84. .311 of consensus"
20780. .22091
/note="LIM4 repeat: matches 1271. .2396 of consensus"
23105. .23475
/note="LTR1B repeat: matches 1. .390 of consensus"
23746. .23852
/note="LIM5A repeat: matches 6185. .6290 of consensus"
24766. .24827
/note="31 copies 2 mer tt 77% conserved"
29279. .29497
/note="AluY repeat: matches 1. .238 of consensus"
29508. .29537
/note="15 copies 2 mer aa 93% conserved"
29705. .29794
/note="L2 repeat: matches 2527. .2612 of consensus"
29997. .30450
/note="match: STS: Em:G28363"
30837. .30899
/note="MIR repeat: matches 102. .164 of consensus"
30910. .31047
/note="AluY/FRAM repeat: matches 177. .310 of consensus"
31280. .31487
/note="MER58A repeat: matches 7. .224 of consensus"
31887. .32107

repeat_region 32535. .33158
/note="AluSg/x repeat: matches 87. .312 of consensus"
34630. .34990
/note="LIM5 repeat: matches 5333. .6002 of consensus"
36142. .36247
/note="THE1C repeat: matches 1. .371 of consensus"
36248. .36509
/note="U6 repeat: matches 1. .107 of consensus"
36977. .37291
/note="LIM2 repeat: matches 6034. .6308 of consensus"
37291. .37353
/note="LIP3 repeat: matches 5853. .6150 of consensus"
37499. .38488
/note="L1 repeat: matches 4746. .4811 of consensus"
38480. .38572
/note="LIM1 repeat: matches 4808. .5794 of consensus"
38553. .39017
/note="LIM1 repeat: matches 4746. .4836 of consensus"
39443. .44686
/note="L1 repeat: matches 4266. .4745 of consensus"
44682. .45457
/note="LIP2 repeat: matches 900. .6146 of consensus"
45682. .45845
/note="L2 repeat: matches 2545. .2710 of consensus"
46265. .46565
/note="AluY repeat: matches 1. .297 of consensus"
47546. .47599
/note="27 copies 2 mer ac 79% conserved"
47552. .47599
/note="12 copies 4 mer acac 81% conserved"
48367. .48595
/note="L2 repeat: matches 1755. .2008 of consensus"
49387. .49709
/note="AluY repeat: matches 1. .312 of consensus"
49767. .50062
/note="AluSg repeat: matches 1. .297 of consensus"
52784. .53168
/note="LTR1A2 repeat: matches 2. .374 of consensus"
56002. .56206
/note="LTR1B repeat: matches 179. .390 of consensus"
56206. .56473
/note="LTR1B repeat: matches 1. .276 of consensus"
56935. .56965
/note="L2 repeat: matches 2672. .2702 of consensus"
57451. .57492
/note="match: STS: Em:G24088"
57691. .57748
/note="21 copies 2 mer aa 88% conserved"
57693. .57748
/note="29 copies 2 mer tg 91% conserved"
57751. .58060
/note="14 copies 4 mer ttgt 91% conserved"
58815. .59127
/note="AluY repeat: matches 1. .294 of consensus"
63084. .63401
/note="AluY repeat: matches 1. .311 of consensus"
63416. .63529
/note="AluSx repeat: matches 1. .310 of consensus"
66317. .66372
/note="MIR repeat: matches 110. .240 of consensus"
66979. .67390
/note="match: STS: Em:G03280"
67989. .68264
/note="AluY repeat: matches 1. .288 of consensus"
68264. .68603
/note="L1 repeat: matches 1. .785 of consensus"
69317. .69372
/note="L1 repeat: matches 1. .785 of consensus"
69317. .69372
/note="14 copies 4 mer acac 85% conserved"
69317. .69372
/note="MIR repeat: matches 1. .418 of consensus"
69317. .69372
/note="AluSx repeat: matches 1. .276 of consensus"
```

Repeat\_region 69790..69889

alignment\_scores:

Quality: 195.00 Length: 33

Ratio: 5.909 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-877-665-4 x AL391478 ..

Align seg 1/1 to: AL391478 from: 1 to: 180625

1 HispHelysProcysaArgAspLysAspLeuAlaTyrCysLeuAsnAspG1 17

54142 CACTCAAAACCTGCGGACAGCAAGACCTGCTGCTCATGATGAG 54191

17 yglucyspheval11legluthtleuthrlyserh1slyshs 33

54192 CGAGTGCTTGATCGAAACCTGACCGGATCCCATAAACACTGTGCG 54240

seq\_name: gb\_ov:AF076618

seq\_documentation\_block:

LOCUS AF076618 3105 bp mRNA linear VRT 18-JUL-1998

DEFINITION Xenopus laevis neuregulin alpha-1 mRNA, complete cds.

ACCESSION AF076618

VERSION AF076618.1 GI:3328216

KEYWORDS African clawed frog.

SOURCE Xenopus laevis

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;

Xenopodinae; Xenopus.

REFERENCE 1 (bases 1 to 3105)

Yang, J.F., Zhou, H., Pun, S., Ip, N.Y., Peng, H.B. and Tsai, K.W.

Cloning of cDNAs encoding xenopus neuregulin: expression in

myotomal muscle during embryo development

Brain Res. Mol. Brain Res. 58 (1-2), 59-73 (1998)

JOURNAL MEDLINE 98352126

REFERENCE 2 (bases 1 to 3105)

Yang, J.F., Zhou, H., Pun, S., Ip, N.Y., Peng, H.B. and Tsai, K.W.

Direct Submission

Submitted (03-JUL-1998) Department of Biology and Biotechnology

Institute, The Hong Kong University of Science and Technology,

Clear Water Bay Road, Hong Kong, China

Location/Qualifiers

1..3105

/organism="Xenopus laevis"

/db\_xref="taxon:8355"

161..2194

/function="acetylcholine receptor-inducing activity"

/note="SP43"

/codon\_start=1

/product="neuregulin alpha-1"

/protein\_id="AAC26804.1"

/db\_xref="GI:3328217"

/translation="MAEKKVKYEGKRGKGRKGRKKAEGSDOGAASPRLKEIKT

OSVOGKRLVLCQAVSEDPKLFKFKGKGEKAKKPKDSEPHIKRIGKSSKSLQ

ISKSSADNGEYKCAVSNQNDGTVNTVNTVIVKPKYNNLLMKYIKYTSVEKSE

STLNLESOKEVYFATKRGDTAGGHLIKGDKPKYCVAGGECYVINGTSSNOF

MCVKPKGTGACRTEDPLRYVRSKHGIEMELQKVLITGICIDLVLVGD

SKNVISSHEVERETETFSSTSHYSTHSTVOTPSHSNGLSESMISEKSYV

ITVSSVENSRTSPSPGRGLNGIGPRDCSTLRARAPDPSRDSERVSAMT

PARSPVPEFKTPIPSKPSCLSTSPSSSLAVSPVAVSPTEERPLIVSPPLRE

KRYDRKTPKTPHNGHNSYHNHNGHSSSLAVSPVAVSPTEERPLIVSPPLRE

LVNSRROKRTKPNHNSYHNHNGHSSSLAVSPVAVSPTEERPLIVSPPLRE

LESASLYHNSRTPSPSTOERLQRLSLANQALCDCKRRMCKTLEFI"

Quality: 126.50 Length: 46

Ratio: 3.614 Gaps: 1

Percent Similarity: 76.087 Percent Identity: 43.478

alignment\_block:

US-09-877-665-4 x AF076618 ..

Align seg 1/1 to: AF076618 from: 1 to: 3105

1 HispHelysProcysaArgAspLysAspLeuAlaTyrCysLeuAsnAspG1 17

722 CACCTTATTAAGTTCACACACAGAAAGACCTGCTGCTCATGAGG 771

17 yglucyspheval11legluthtleuthrlyserh1slyshs...Cysa 33

772 AGAGTGTATGATGACGACGATATACAGACAGCAACGATTCATGCA 821

33 rgcyslysglulgltyrtinglglvalargcysaspGln 45

822 AGTGCACACCTGATGCTGACGAGCAAGATGACATGCA 859

seq\_name: gb\_pat:AR100663

seq\_documentation\_block:

LOCUS AR100663 156 bp DNA linear PAT 14-FEB-2001

DEFINITION Sequence 2 from patent US 6080845.

ACCESSION AR100663

VERSION AR100663.1 GI:12811111

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 156)

Carahan, J.F.

Monoclonal antibody against utricular epithelium

Patent: US 6080845-A 2 27-JUN-2000;

location/Qualifiers

1..156

/organism="unknown"

BASE COUNT 47 a 31 c 42 g 36 t

ORIGIN

alignment\_scores:

Quality: 116.50 Length: 48

Ratio: 3.149 Gaps: 1

Percent Similarity: 77.083 Percent Identity: 33.333

alignment\_block:

US-09-877-665-4 x AR100663 ..

Align seg 1/1 to: AR100663 from: 1 to: 156

1 HispHelysProcysaArgAspLysAspLeuAlaTyrCysLeuAsnAspG1 17

4 CATCTTATAAATGTCGCGAAGAGAAACTTCTGCTGATGAGG 53

17 yglucyspheval11legluthtleuthrlyserh1slyshs...Cysa 33

54 GGAATGCTTATGCTGAGAAAGACCTTCAACCCCTCGAATCTTGCA 103

33 rgcyslysglulgltyrtinglglvalargcysaspGln 47

104 AGTGCACACCTGATGCTGACGAGCAAGATGCAAACTACGTA 147

seq\_name: gb\_pat:AR022501

seq\_documentation\_block:

LOCUS AR022501 192 bp DNA linear PAT 05-DEC-1998

DEFINITION Sequence 155 from patent US 5792849.

ACCESSION AR022501

VERSION AR022501.1 GI:3976563

KEYWORDS



```

BASE COUNT          61 a      39 c      50 g      42 t
ORIGIN              /organism="unknown"

alignment_scores:
    Quality: 113.50           Length: 46
    Ratio: 3.338             Gaps: 1
    Percent Similarity: 73.913   Percent Identity: 34.783

alignment_block:
US-09-877-665-4 x AR135006 ..

Align seg 1/1 to: AR135006 from: 1 to: 192

1 HspshelysProCysArGAsPlysaSPleuaLaTyrcYseLsuhsnaSpcl 17
||||| ||| ::|||:: :|||::||| |
4 CATCTTGTCAGAGTGTGCAGAGAAGAAAACCTTCTGTGTGAATGAGG 53

17 yglucYsPheValIleGluthrLeuthrclgYserHlsYhs...Cysa 33
||||| ||| ::|||:: :|||::||| |
54 CGAGTCTTCATGGTGAAGAAGACCTTCAAAATCCCTCAAGATACTGTGCA 103

33 rgcYslYsglUglyTYrgInglYvalarGcYsaSpgin 45
::::: |||::: |||::: |||:::
104 AGTCCAACCTCGATTCACTGAGCAGCATGTACTGAG 141

seq_name: gb_pat:AR141835

seq_documentation_block:
LOCUS       AR141835               192 bp    DNA        linear     PAT 08-AUG-2001
DEFINITION  Sequence 155 from patent US 6147190.
ACCESSION   AR141835
VERSION     AR141835.1 GI:15101351
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 192)
AUTHORS    Goodale,L.A., Stroobant,P., Minghettil,L., Waterfield,M.,
            Marchionni,M., Chen,M.Su and Hiles,I.
TITLE      Glial mitogenic factors: their preparation and use
JOURNAL    Patent: US 6147190-A 155 14-NOV-2000;
FEATURES
            source
                location/qualifiers
                    1..192
                        /organism="unknown"
BASE COUNT      61 a      39 c      50 g      42 t
ORIGIN

```

```

LOCUS       ARI143352                192 bp      DNA          linear      PAT 08-AUG-2001
DEFINITION   Sequence 155 from patent US 6204241.
ACCESSION    ARI143352
VERSION      ARI143352.1  GI:15104638
KEYWORDS
SOURCE       Unknown.
ORGANISM     Unknown.
REFERENCE    1 (bases 1 to 192)
AUTHORS      Goodearl,A.,David, Stroobant,P., Minghetti,L., Waterfield,M.,
              Matchionni,M., Chen,M.Su and Hiles,I.
TITLE        Method for treating nervous system pathophysologies using glial
              growth factors
              Patent: US 6204241-A 155 20-MAR-2001;
              Location/Qualifiers
                source          1..192
                /organism="unknown"
BASE COUNT   61 a      39 c      50 g      42 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985

```







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 08:42:13 ; Search time 13.02 Seconds

(without alignments)  
(8.172 Million cell updates/sec)

Title: US-09-877-665-4

Sequence: 1 HFKRCRDLAYCLNDGEFCF.....SHKRCRKEGVRCDOFL 47

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

231628 segs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Database :

1: Issued Patents.AA:\*  
2: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/BACKFILES1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	277	100.0	47	3	US-08-899-437-4
2	277	100.0	47	3	US-08-899-437-8
3	277	100.0	47	4	US-08-126-121-4
4	277	100.0	47	4	US-09-126-121-4
5	277	100.0	47	4	US-09-126-121-8
6	277	100.0	360	3	US-08-899-437-7
7	277	100.0	360	4	US-09-126-121-7
8	277	100.0	362	3	US-08-899-437-3
9	277	100.0	362	4	US-09-126-121-3
10	277	100.0	696	3	US-08-899-437-23
11	277	100.0	713	3	US-09-126-121-23
12	277	100.0	713	4	US-08-899-437-2
13	277	100.0	720	3	US-09-126-121-2
14	277	100.0	720	4	US-08-899-437-6
15	116.5	42.1	52	1	US-08-417-640A-1
16	116.5	42.1	52	1	US-08-760-815-1
17	116.5	42.1	52	2	US-08-761-038-1
18	116.5	42.1	52	3	US-09-238-182-1
19	113.5	41.0	49	3	US-08-899-437-14
20	113.5	41.0	49	4	US-09-126-121-14
21	113.5	41.0	50	3	US-08-753-007A-12
22	113.5	41.0	50	4	US-09-398-496-12
23	113.5	41.0	52	1	US-08-417-640A-3
24	113.5	41.0	52	1	US-08-760-815-3
25	113.5	41.0	52	2	US-08-761-038-3
26	113.5	41.0	54	1	US-08-179-481-111
27	113.5	41.0	63	3	US-08-341-018-62

28	113.5	41.0	63	4	US-08-470-335-221	Sequence 221, App
29	113.5	41.0	63	4	US-08-470-339-221	Sequence 221, App
30	113.5	41.0	66	1	US-07-847-743B-10	Sequence 10, App
31	113.5	41.0	66	1	US-08-456-201-10	Sequence 10, App
32	113.5	41.0	66	2	US-08-456-241-10	Sequence 10, App
33	113.5	41.0	66	4	US-09-020-880-2	Sequence 2, App1
34	113.5	41.0	66	5	PCT-US92-04295A-10	Sequence 10, App1
35	113.5	41.0	83	3	US-08-341-018-70	Sequence 225, App
36	113.5	41.0	83	4	US-08-470-335-225	Sequence 225, App
37	113.5	41.0	88	3	US-08-470-339-225	Sequence 225, App
38	113.5	41.0	88	3	US-08-341-018-68	Sequence 68, App1
39	113.5	41.0	88	4	US-08-470-335-224	Sequence 224, App
40	113.5	41.0	88	4	US-08-470-339-224	Sequence 224, App
41	113.5	41.0	95	1	US-07-847-743B-14	Sequence 14, App1
42	113.5	41.0	95	1	US-08-456-201-14	Sequence 14, App1
43	113.5	41.0	95	2	US-08-330-161-12	Sequence 12, App1
44	113.5	41.0	95	2	US-08-456-241-14	Sequence 12, App1
45	113.5	41.0	95	2	US-08-440-401-12	Sequence 12, App1

#### ALIGNMENTS

RESULT 1  
US-08-899-437-4  
Sequence 4, Application US/08899437  
Patent No. 6121415  
GENERAL INFORMATION:  
APPLICANT: Gilead  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESSES:  
ADDRESS: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/899,437  
FILING DATE: 24-Jul-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Delidre L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 47 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: NRG3 EGF-like domain/amino acid seq.  
LOCATION: 1-47  
IDENTIFICATION METHOD:  
OTHER INFORMATION:

US-08-899-437-4

Query Match 100.0%; Score 277; DB 3; Length 47;  
Best Local Similarity 100.0%; Pred. No. 1.3e-26;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFKRCRDLAYCLNDGEFCFVETLTGSHKRCRKEGVRCDOFL 47

DB 1 HFKPCRDILAYCLNDEGCFYIEITLGSNHCRCRCKEGYGVRCDOFL 47

## RESULT 2

US-08-899-437-8 Application US/08899437

Sequence 8 6121415

Patent No. 6121415

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao

TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/899,437

FILING DATE: 24-Jul-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Conley, Delidre L.

REGISTRATION NUMBER: 36,487

REFERENCE/DOCKET NUMBER: P1084R1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/952-2066

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 47 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

FEATURE:

NAME/KEY: NR3 EGF-like domain/amino acid seq.

LOCATION: 1-47

IDENTIFICATION METHOD:

OTHER INFORMATION:

US-08-899-437-8

Query Match 100.0%; Score 277; DB 3; Length 47;

Best Local Similarity 100.0%; Pred. No. 1.3e-26;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 HFKPCRDILAYCLNDEGCFYIEITLGSNHCRCRCKEGYGVRCDOFL 47

RESULT 3

US-09-126-121-4

Sequence 4 Application US/09126121

Patent No. 6252051

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao

TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/126,121

FILING DATE: 30-Jul-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Conley, Delidre L.

REGISTRATION NUMBER: 36,487

REFERENCE/DOCKET NUMBER: P1084R1D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/952-2066

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 47 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

FEATURE:

NAME/KEY: NR3 EGF-like domain/amino acid seq.

LOCATION: 1-47

IDENTIFICATION METHOD:

OTHER INFORMATION:

US-09-126-121-4

Query Match 100.0%; Score 277; DB 4; Length 47;

Best Local Similarity 100.0%; Pred. No. 1.3e-26;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 HFKPCRDILAYCLNDEGCFYIEITLGSNHCRCRCKEGYGVRCDOFL 47

RESULT 4

US-09-126-121-8

Sequence 8 Application US/09126121

Patent No. 6252051

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao

TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/126,121

FILING DATE: 30-Jul-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Conley, Delidre L.

REGISTRATION NUMBER: 36,487

REFERENCE/DOCKET NUMBER: P1084R1D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/952-2066

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 47 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

FEATURE:

NAME/KEY: NR3 EGF-like domain/amino acid seq.

LOCATION: 1-47

IDENTIFICATION METHOD:

OTHER INFORMATION:

US-09-126-121-8

Query Match 100.0%; Score 277; DB 4; Length 47;

Best Local Similarity 100.0%; Pred. No. 1.3e-26;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 HFKPCRDILAYCLNDEGCFYIEITLGSNHCRCRCKEGYGVRCDOFL 47

RESULT 5

US-09-126-121-4

Sequence 4 Application US/09126121

Patent No. 6252051

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao

TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/126,121

FILING DATE: 30-Jul-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Conley, Delidre L.

REGISTRATION NUMBER: 36,487

REFERENCE/DOCKET NUMBER: P1084R1D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/952-2066

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 47 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

FEATURE:

NAME/KEY: NR3 EGF-like domain/amino acid seq.

LOCATION: 1-47

IDENTIFICATION METHOD:

OTHER INFORMATION:

US-09-126-121-4

Query Match 100.0%; Score 277; DB 4; Length 47;

Best Local Similarity 100.0%; Pred. No. 1.3e-26;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 HFKPCRDILAYCLNDEGCFYIEITLGSNHCRCRCKEGYGVRCDOFL 47

RESULT 6

US-09-126-121-4

Sequence 4 Application US/09126121

Patent No. 6252051

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao

TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/126,121

FILING DATE: 30-Jul-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Conley, Delidre L.

REGISTRATION NUMBER: 36,487

REFERENCE/DOCKET NUMBER: P1084R1D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/952-2066

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 47 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

FEATURE:

NAME/KEY: NR3 EGF-like domain/amino acid seq.

LOCATION: 1-47

IDENTIFICATION METHOD:

OTHER INFORMATION:

US-09-126-121-4

Query Match 100.0%; Score 277; DB 4; Length 47;

Best Local Similarity 100.0%; Pred. No. 1.3e-26;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 HFKPCRDILAYCLNDEGCFYIEITLGSNHCRCRCKEGYGVRCDOFL 47

RESULT 7

US-09-126-121-4

Sequence 4 Application US/09126121

Patent No. 6252051

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao

TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/126,121

FILING DATE: 30-Jul-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Conley, Delidre L.

REGISTRATION NUMBER: 36,487

REFERENCE/DOCKET NUMBER: P1084R1D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/952-2066

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 47 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

FEATURE:

NAME/KEY: NR3 EGF-like domain/amino acid seq.

LOCATION: 1-47

IDENTIFICATION METHOD:

OTHER INFORMATION:

US-09-126-121-4

Query Match 100.0%; Score 277; DB 4; Length 47;

Best Local Similarity 100.0%; Pred. No. 1.3e-26;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 HFKPCRDILAYCLNDEGCFYIEITLGSNHCRCRCKEGYGVRCDOFL 47

RESULT 8

US-09-126-121-4

Sequence 4 Application US/09126121

Patent No. 6252051

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao

TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

LENGTH: 47 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: NR3 EGF-like domain/amino acid seq.  
LOCATION: 1-47  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-09-126-121-8

Query Match 100.0%; Score 277; DB 4; Length 47;  
Best Local Similarity 100.0%; Pred. No. 1.3e-26;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFKPCRDKLAYCLNDGCEFYETLTGSHKRCRKEGYGVRCDOFL 47  
DB 1 HFKPCRDKLAYCLNDGCEFYETLTGSHKRCRKEGYGVRCDOFL 47

RESULT 5  
US-08-899-437-7  
Sequence 7, Application US/08899437  
Patent No. 6121415  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
TITLE OF INVENTION: Ligands and Uses Therefor  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/899,437  
FILING DATE: 24-Jul-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Delidre L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 360 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: hNRG3 extracellular domain/Amino Acidseq  
LOCATION: 1-360  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-08-899-437-7

Query Match 100.0%; Score 277; DB 3; Length 360;  
Best Local Similarity 100.0%; Pred. No. 1e-25;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFKPCRDKLAYCLNDGCEFYETLTGSHKRCRKEGYGVRCDOFL 47  
DB 286 HFKPCRDKLAYCLNDGCEFYETLTGSHKRCRKEGYGVRCDOFL 332

RESULT 6  
US-09-126-121-7  
Sequence 7, Application US/09126121  
Patent No. 6252051  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
TITLE OF INVENTION: Ligands and Uses Therefor  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/126,121  
FILING DATE: 30-Jul-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Delidre L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 360 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: hNRG3 extracellular domain/Amino Acidseq  
LOCATION: 1-360  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-09-126-121-7

Query Match 100.0%; Score 277; DB 4; Length 360;  
Best Local Similarity 100.0%; Pred. No. 1e-25;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFKPCRDKLAYCLNDGCEFYETLTGSHKRCRKEGYGVRCDOFL 47  
DB 286 HFKPCRDKLAYCLNDGCEFYETLTGSHKRCRKEGYGVRCDOFL 332

RESULT 7  
US-08-899-437-3  
Sequence 3, Application US/08899437  
Patent No. 6121415  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
TITLE OF INVENTION: Ligands and Uses Therefor  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,437
FILING DATE: 24-Jul-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Delidre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 362 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
FEATURE:
NAME/KEY: MNRG3 extracellular domain amino acid seq
LOCATION: 1-362
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-08-899-437-3

Query Match      100.0%; Score 277; DB 3; Length 362;
Best Local Similarity 100.0%; Pred. No. 1e-25;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDLDAYCLNDGECFVETLTGSHKRCRCKEGYGVRCDOFL 47
DB 288 HFKPCRDLDAYCLNDGECFVETLTGSHKRCRCKEGYGVRCDOFL 334

RESULT 8
US-09-126-121-3
Sequence 3, Application US/09126121
Patent No. 6252051
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
TITLE OF INVENTION: Ligands and Uses Therefor
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/126,121
FILING DATE: 30-Jul-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Delidre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 362 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
```

```

FEATURE:
NAME/KEY: MNRG3 extracellular domain amino acid seq
LOCATION: 1-362
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-09-126-121-3

Query Match      100.0%; Score 277; DB 4; Length 362;
Best Local Similarity 100.0%; Pred. No. 1e-25;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDLDAYCLNDGECFVETLTGSHKRCRCKEGYGVRCDOFL 47
DB 288 HFKPCRDLDAYCLNDGECFVETLTGSHKRCRCKEGYGVRCDOFL 334

RESULT 9
US-08-899-437-23
Sequence 23, Application US/08899437
Patent No. 6121415
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
TITLE OF INVENTION: Ligands and Uses Therefor
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,437
FILING DATE: 24-Jul-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Delidre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 696 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
FEATURE:
NAME/KEY: Human MNRG3B2
LOCATION: 1-696
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-08-899-437-23

Query Match      100.0%; Score 277; DB 3; Length 696;
Best Local Similarity 100.0%; Pred. No. 2e-25;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDLDAYCLNDGECFVETLTGSHKRCRCKEGYGVRCDOFL 47
DB 286 HFKPCRDLDAYCLNDGECFVETLTGSHKRCRCKEGYGVRCDOFL 332

RESULT 10
US-09-126-121-23
```

Sequence 23, Application US/09126121  
Patent No. 6252051  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/126,121  
FILING DATE: 30-Jul-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Delidre L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 696 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: Human NRG3B2  
LOCATION: 1-696  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-09-126-121-23

Query Match 100.0%; Score 277; DB 4; Length 696;  
Best Local Similarity 100.0%; Pred. No. 2e-25;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDKLATCLNDGECFVETLTGSHKCRCKRGYGVRCDOFL 47  
|||||  
Db 286 HFKPCRDKLATCLNDGECFVETLTGSHKCRCKRGYGVRCDOFL 332

RESULT 11  
US-08-899-437-2  
Sequence 2, Application US/08899437  
Patent No. 6121415  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/899,437  
FILING DATE: 24-Jul-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Delidre L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 713 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: Mouse NRG3 (mNRG3)/amino acid seq.  
LOCATION: 1-713  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-08-899-437-2

Query Match 100.0%; Score 277; DB 3; Length 713;  
Best Local Similarity 100.0%; Pred. No. 2.1e-25;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDKLATCLNDGECFVETLTGSHKCRCKRGYGVRCDOFL 47  
|||||  
Db 288 HFKPCRDKLATCLNDGECFVETLTGSHKCRCKRGYGVRCDOFL 334

RESULT 12  
US-09-126-121-2  
Sequence 2, Application US/09126121  
Patent No. 6252051  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/126,121  
FILING DATE: 30-Jul-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Delidre L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 713 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: Mouse NRG3 (mNRG3)/amino acid seq.  
LOCATION: 1-713

IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-09-126-121-2

Query Match 100.0%; Score 277; DB 4; Length 713;  
Best Local Similarity 100.0%; Pred. No. 2.1e-25;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDRLAYCLNDGECFVETLTGSHKCRCKEGYGVRCDOFL 47  
DB 288 HFKPCRDRLAYCLNDGECFVETLTGSHKCRCKEGYGVRCDOFL 334

RESULT 13

US-08-899-437-6  
Sequence 6, Application US/08899437  
Patent No. 6121415

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:

ADDRESSER: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Winpatin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/899,437

FILING DATE: 24-Jul-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Conley, Delidre L.

REGISTRATION NUMBER: 36,487

REFERENCE/DOCKET NUMBER: P1084R1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-2066

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 720 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

FEATURE:

NAME/KEY: hNRG3B1 amino acid sequence

LOCATION: 1-720

IDENTIFICATION METHOD:

OTHER INFORMATION:

US-08-899-437-6

Query Match 100.0%; Score 277; DB 3; Length 720;  
Best Local Similarity 100.0%; Pred. No. 2.1e-25;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDRLAYCLNDGECFVETLTGSHKCRCKEGYGVRCDOFL 47  
DB 286 HFKPCRDRLAYCLNDGECFVETLTGSHKCRCKEGYGVRCDOFL 332

RESULT 14

US-09-126-121-6  
Sequence 6, Application US/09126121  
Patent No. 6252051  
GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:

ADDRESSER: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Winpatin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/126,121

FILING DATE: 30-Jul-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Conley, Delidre L.

REGISTRATION NUMBER: 36,487

REFERENCE/DOCKET NUMBER: P1084R1D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-2066

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 720 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

FEATURE:

NAME/KEY: hNRG3B1 amino acid sequence

LOCATION: 1-720

IDENTIFICATION METHOD:

OTHER INFORMATION:

US-09-126-121-6

Query Match 100.0%; Score 277; DB 4; Length 720;  
Best Local Similarity 100.0%; Pred. No. 2.1e-25;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDRLAYCLNDGECFVETLTGSHKCRCKEGYGVRCDOFL 47  
DB 286 HFKPCRDRLAYCLNDGECFVETLTGSHKCRCKEGYGVRCDOFL 332

RESULT 15

US-08-417-640A-1

Sequence 1, Application US/08417640A

Patent No. 5670342

GENERAL INFORMATION:

APPLICANT: Carnahan, Josette F.

APPLICANT: Hara, Shinichi

APPLICANT: Lu, Hsiang S.

APPLICANT: Mayer, John P.

APPLICANT: Yoshinaga, Steven K.

TITLE OF INVENTION: NDF Peptides

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSER: Amgen Inc.

STREET: 1840 Delnavilland Drive

CITY: Thousand Oaks

STATE: California

COUNTRY: USA

ZIP: 91320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 08:49:33 ; Search time 25.6 Seconds  
(without alignments)  
317.608 Million cell updates/sec

Title: US-09-877-665-4  
Sequence: 1 HFKPCRDLDALCYLNDGECF.....SHKHCKCKEGYGVKCDPL 47

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea:\*  
2: SP bacteria:\*  
3: SP fungi:\*  
4: SP human:\*  
5: SP\_invertebrate:\*  
6: SP\_mammal:\*  
7: SP\_mhc:\*  
8: SP\_organelle:\*  
9: SP\_phage:\*  
10: SP\_plant:\*  
11: SP\_rodent:\*  
12: SP\_virus:\*  
13: SP\_vertebrate:\*  
14: SP\_unclassified:\*  
15: SP\_virus:\*  
16: SP\_bacteriap:\*  
17: SP\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113.5	41.0	461	11	035947 mesocricetu
2	111.5	40.3	298	11	09ESa9 rattus norv
3	111.5	40.3	695	11	09ESB0 rattus norv
4	104.5	37.7	241	6	007112 bos taurus
5	104.5	37.7	296	4	0961B3
6	102.5	37.0	111	11	09ESa8 rattus norv
7	102.5	37.0	136	11	09ESa7 rattus norv
8	102.5	37.0	156	11	09ESa6 rattus norv
9	102.5	37.0	317	11	09ESa3 rattus norv
10	102.5	37.0	323	11	09ESa2 rattus norv
11	102.5	37.0	342	11	09ESa1 rattus norv
12	102.5	37.0	700	11	09ESb1 rattus norv
13	102.5	37.0	782	11	09ESa5 rattus norv
14	99	35.7	89	12	091M20
15	92.5	33.4	2192	5	001768
16	91.5	33.0	1213	13	090Y54

17	89	32.1	162	11	0920L5	0920L5 rattus norv
18	89	32.1	1241	4	09UKK5	09UKK5 homo sapien
19	89	32.1	1241	4	09UKD4	09UKD4 homo sapien
20	89	32.1	1241	4	09B2V3	09B2V3 homo sapien
21	88	31.8	1239	11	P706Z8	P706Z8 rattus norv
22	87	31.4	217	5	09VVT6	09VVT6 dirosophila
23	87	31.4	597	11	035727	035727 mus muscicu
24	85.5	30.9	1193	13	090819	090819 gallus gall
25	85.5	30.9	1214	13	090YD2	090YD2 xenopus lae
26	85.5	30.9	1218	4	014902	014902 homo sapien
27	85.5	30.9	1218	4	015122	015122 homo sapien
28	85.5	30.9	1218	4	015816	015816 homo sapien
29	85.5	30.9	1227	4	P78504	P78504 homo sapien
30	84.5	30.5	177	11	09JUM4	09JUM4 rattus norv
31	84.5	30.5	1218	11	09QXX0	09QXX0 mus muscicu
32	84.5	30.5	1219	11	063722	063722 rattus norv
33	83.5	30.1	127	12	0909F3	0909F3 ectomella
34	83	30.0	162	11	061521	061521 mus muscicu
35	83	30.0	861	11	09QW58	09QW58 mus sp. mot
36	83	30.0	862	11	099JC2	099JC2 mus muscicu
37	82.5	29.8	445	5	09W3W5	09W3W5 dirosophila
38	82.5	29.8	456	5	0961F3	0961F3 dirosophila
39	82	29.6	264	5	020559	020559 caenorhabd
40	81	29.2	78	12	091T36	091T36 lumpy skin
41	80.5	29.1	125	12	041504	041504 compox viru
42	80.5	29.1	178	4	096F48	096F48 homo sapien
43	80.5	29.1	260	4	095898	095898 homo sapien
44	80.5	29.1	530	4	096CM9	096CM9 homo sapien
45	80.5	29.1	1242	13	090Y57	090Y57 brachydanto

## ALIGNMENTS

RESULT	1	PRELIMINARY	PRT	461 AA.
035947	AC	035947		
ID	035947	035947		
DT	01-JAN-1998 (Tremblrel. 05, Created)			
DT	01-JAN-1998 (Tremblrel. 05, Last sequence update)			
DT	01-JUN-2001 (Tremblrel. 17, Last annotation update)			
DE	PRO-NEUREGULIN-1, ISOFORM ALPHA 2B PRECURSOR.			
GN	NRG1 OR NDF.			
OS	Mesocricetus auratus (Golden hamster).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;			
OC	Mesocricetus.			
OX	NCBI_Taxid=10036;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM ALPHA2B), AND SEQUENCE OF 64-81.			
RC	TISSUE=EMBRYO;			
RX	MEDLINE-98196966; PubMed-9537646;			
RA	Velasco J.A., Feljoo E., Avila M.A., Notario V.;			
RT	Secretion of new differentiation factor-like polypeptides by cph-			
RT	transformed fibroblasts: cloning and characterization of Syrian			
RT	hamster neurogenin cDNAs.;			
RI	Mol. Cell Biol. 21:156-163(1998).			
RL	Mol. Cell Biol. 21:156-163(1998).			
CC	-1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE			
CC	RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,			
CC	RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND			
CC	ACTIVATION OF THE ERBB RECEPTORS. MAY PLAY AN IMPORTANT ROLE IN			
CC	PROVIDING GROWTH ADVANTAGE IN NEOPLASTIC CELLS.			
CC	-1- SUBUNIT: THE CYTOPLASMIC DOMAIN INTERACTS WITH THE LIM DOMAIN			
CC	REGION OF LIM1 (BY SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: EXISTS AS TYPE I MEMBRANE PROTEIN AND AS A			
CC	PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-			
CC	BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).			
CC	-1- ALTERNATIVE PRODUCTS: DIFFERENT ISOFORMS ARE PRODUCED BY			
CC	ALTERNATIVE SPLICING. THE SEQUENCE SHOWN IS THAT OF ISOFORM			
CC	ALPHA2B/CLONE PM3.			
CC	-1- TISSUE SPECIFICITY: EXPRESSED AT HIGHER LEVEL AFTER NEOPLASMIC			
CC	TRANSFORMATION OF CELLS.			
CC	-1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION			

```

CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
CC DIMERIZATION (BY SIMILARITY).
CC -1- DOMAIN: EREB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
CC DOMAIN.
CC -1- PFM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
CC FORM (BY SIMILARITY).
CC -1- PFM: EXTENSIVE GLYCOSYLATION PRECEDES PROTEOLYTIC CLEAVAGE (BY
CC SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
DR EMBL: U96612; AAB71812.1; -.
DR HSSP: Q12784; IHR.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR003598; Ig-C2.
DR InterPro: IPR003006; Ig-MHC.
DR InterPro: IPR002154; Neuregulin.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00047; Ig; 1.
DR Pfam: PF02158; Neuregulin; 1.
DR PRINTS: PRO1089; NEUREGULIN.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00408; IGC2; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 1.
DR Growth factor: EGF-like domain; Immunoglobulin domain; Glycoprotein;
DR Transmembrane; Alternative splicing.
FT PROPEP 1 13
FT CHAIN 14 461
FT CHAIN 14 241
FT DOMAIN 14 242
FT TRANSSEM 243 265
FT DOMAIN 266 461
FT DOMAIN 50 119
FT DOMAIN 165 177
FT DOMAIN 178 222
FT DISULFID 57 112
FT DISULFID 182 196
FT DISULFID 190 210
FT DISULFID 212 221
FT CARBOHYD 73 73
FT CARBOHYD 120 120
FT CARBOHYD 126 126
FT CARBOHYD 164 164
SQ SEQUENCE 461 AA; 50890 MW; 935C9560F7148336 CRC64;

Query Match 41.0%; Score 113.5; DB 11; Length 461;
Best Local Similarity 34.8%; Pred. No. 1e-07;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

QY 1 HFKPCRDLDAYCLNDGECFVIEITLGSKH-CRCKEGYGVRCDO 45
Db 178 HIKCAKEKTEFCVNGGECFTVKDLSNPSRYLCKCPGFTGARCTE 223

RESULT 2
Q9ESA9 PRELIMINARY; PRT; 298 AA.
AC Q9ESA9;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE SMDF NEUREGULIN ALPHA 2B (FRAGMENT).
GN NRGL
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-BDIX;
RA Carroll S.L., Anderson K.D., Frohner P.W.;
RT "Structural and Functional Diversity of SMDF Neuregulin Splice
RT Variants Expressed in the Adult Rat Nervous System.";
RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF194440; AAC28429.1; -.
DR HSSP: Q12784; IHR.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002154; Neuregulin.
DR InterPro: IPR002114; PTS_HPR_ser.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF02158; Neuregulin; 1.
DR PRINTS: PRO1089; NEUREGULIN.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00001; EGF-like; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS00589; PTS_HPR_ser; UNKNOWN_1.
FT NON_TER 1 1
FT NON_TER 298 298
SQ SEQUENCE 298 AA; 32851 MW; BD76F014C2B33026 CNC64;

Query Match 40.3%; Score 111.5; DB 11; Length 298;
Best Local Similarity 34.8%; Pred. No. 1.3e-07;
Matches 16; Conservative 13; Mismatches 16; Indels 1; Gaps 1;

QY 1 HFKPCRDLDAYCLNDGECFVIEITLGSKH-CRCKEGYGVRCDO 45
Db 48 HIKCAKEKTEFCVNGGECFTVKDLSNPSRYLCKCPGFTGARCTE 93

RESULT 3
Q9ESB0 PRELIMINARY; PRT; 695 AA.
AC Q9ESB0;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE SMDF NEUREGULIN ALPHA 2A.
GN NRGL
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BDIX;
RA Carroll S.L., Anderson K.D., Frohner P.W.;
RT "Structural and Functional Diversity of SMDF Neuregulin Splice
RT Variants Expressed in the Adult Rat Nervous System.";
RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF194439; AAC28428.1; -.
DR HSSP: Q12784; IHR.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002154; Neuregulin.
DR InterPro: IPR002114; PTS_HPR_ser.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF02158; Neuregulin; 1.
DR PRINTS: PRO1089; NEUREGULIN.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00001; EGF-like; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS00589; PTS_HPR_ser; UNKNOWN_1.
SQ SEQUENCE 695 AA; 75646 MW; 5277F20BA2FB6878 CRC64;

Query Match 40.3%; Score 111.5; DB 11; Length 695;
Best Local Similarity 34.8%; Pred. No. 3.1e-07;
Matches 16; Conservative 13; Mismatches 16; Indels 1; Gaps 1;

QY 1 HFKPCRDLDAYCLNDGECFVIEITLGSKH-CRCKEGYGVRCDO 45
Db 234 HIKCAKEKTEFCVNGGECFTVKDLSNPSRYLCKCPGFTGARCTE 279

```

	RESULT	4
ID	007112	PRELIMINARY; PTG; 241 AA.
AC	007112;	
DT	01-JAN-1998 (TREMBREL .05, Created)	
DT	01-JUN-1998 (TREMBREL .05, last sequence update)	
DE	NEUREGULIN-1, GLIAL GROWTH FACTOR 5 ISOFORM PRECURSOR (GFBBP5).	
GN	NNG1 OR GGF.	
OS	Bos taurus (Bovine).	
OC	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
CC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;	
OX	Bovidae; Bos.	
RN	NCBI_TaxID=9913;	
RP	SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.	
RC	TISSUE=POSTERIOR PITUITARY;	
RX	MEDLINE=93205115; PubMed=8096067;	
RA	Marchionni M.A., Gooden A.D.J., Chen M.S., Bermingham-McDonogh O.,	
RA	Kirk C., Hendricks M., Denehy F., Mitsumi D., Suchalter J.,	
RA	Kobayashi K., Wroblewski D., Lynch C., Baldaire M., Hiles I.,	
RA	Davis J.B., Hsuan J.-T., Totty N.E., Otsu M., McBurney R.N.,	
RT	Waterfield M.D., Stroobant P., Gwynne D.;	
RT	"Glial growth factors are alternatively spliced erbB2 ligands	
RL	expressed in the nervous system.";	
CC	Nature 362:312-318(1993).	
CC	- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE	
CC	RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,	
CC	RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND	
CC	ACTIVATION OF THE ERBB RECEPTORS (BY SIMILARITY).	
CC	- SUBCELLULAR LOCATION: SECRETED (PROBABLE).	
CC	- TISSUE SPECIFICITY: EXPRESSED IN NERVOUS SYSTEM: SPINAL CORD AND	
CC	BRAIN.	
CC	- ALTERNATIVE PRODUCTS: AT LEAST FIVE DIFFERENT ISOFORMS; GFBBP1,	
CC	GFBBP2, GFBBP3, GFBBP4 AND GFBBP5 (SHOWN HERE); ARE PRODUCED	
CC	BY ALTERNATIVE SPLICING.	
CC	- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.	
CC	- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.	
CC	- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.	
DR	EMBL: L12259; AAA30540.1; -.	
DR	HSSP: Q12784; IHRE.	
DR	InterPro: IPRO00561; EGF-like.	
DR	InterPro: IPRO03598; IG_c2.	
DR	InterPro: IPRO03006; IG_MHC.	
DR	Pfam: PF00008; EGF_1.	
DR	Pfam: PF00047; Ig_1.	
DR	pSMART; SM00181; EGF; 1.	
DR	pSMART; SM00408; IGc2; 1.	
DR	PROSITE; PS00022; EGF_1; UNKNOWN_1.	
DR	PROSITE; PS01862; EGF_2; FALSE_NEG.	
DR	Growth factor: EGF-like domain; Immunoglobulin domain;	
DR	Alternative splicing	
FT	PROPEP	1 19
FT	CHAIN	20 241
FT	DOMAIN	50 117
FT	DOMAIN	165 179
FT	DOMAIN	178 222
FT	DISULFID	182 196
FT	DISULFID	190 210
FT	DISULFID	212 221
FT	VARSPLIC	134 156
FT	VARSPLIC	157 241
FT	VARSPIC	241 AA; 25955 MW; BF57129/EBDA9796 CXC64;
FT	SEQUENCE	

Query Match	37.78;	Score 104.5;	DB 6;	Length 241;
Best Local Similarity	31.28;	Pred. No. 9.7e-07;		
Matches 15;	Conservative 14;	Mismatches 18;	Indels 1;	Gaps 1

QY 1 HFKPCRDKDLAYCLNDGECFVIEITLTGSHKH-CRCKEGYQGVRCDDFL 47  
 178 HLVKCAEKEKTFVNGGECFEMVKDLSNPSRYLCKCPNEFTTGDRQONY 225

RESULT	5		
096IB3			
ID	096IB3	PRELIMINARY;	PRT. 296 AA.
AC	096IB3		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	NEUROGULIN 1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxId=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=BRAIN, AND NEUROBLASTOMA;		
RA	Strausberg R.;		
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC007675; AA07675.1; -.		
SQ	SEQUENCE 296 AA; 31699 MW; 0F5C48C8465D6649 CRC64;		

Query Match 37.7%, Score 104.5; DB 4; Length 296;  
 Best Local Similarity 31.2%; Pred. No. 1.2e-06;  
 Matches 15; Conservative 14; Mismatches 18; Indels 1; Gaps 1

Q	SEQUENCE	111 AA	1219 MW	CCBBB870584C9F8C CRC64
Q	NON_TER	111	111	
FT	NON_TER	1	1	
DR	PROSITE: PS00014; ER_TARGET; UNKNOWN_1.			
DR	SMART: SM00001; EGF_like; 1.			
DR	SMART: SM00181; EGF; 1.			
DR	Pfam: PF00008; EGF; 1.			
DR	InterPro: IPR000865; ER_target.			
DR	InterPro: IPR000561; EGF-like.			
DR	HSSP: Q12784; IHKE.			
DR	EMBL: AB194441; AAG28430.1; -.			
RT	Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.			
RT	Varients Expressed in the Adult Rat Nervous System";			
RC	Structural and Functional Diversity of SMDF Neuregulin Splice			
RC	Carroll S.L., Anderson K.D., Frohner P.W.;			
RC	TISSE-XXOMIZED LUMBAR DORSAL ROOT GANGLION/SPINAL CORD.			
RC	SPRAIN-SPRAGUE-DAWLEY;			
RP	SEQUENCE FROM N.A.			
RP	[[1]]			
OC	NCBI_TaxID=10116;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus			
OS	Rattus norvegicus (Rat).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OT	01-MAR-2001 (TEMBREL. 16, Last sequence update)			
OT	01-MAR-2001 (TEMBREL. 16, Last sequence update)			
DE	SMDF NEUREGULIN BETA 2 (FRAGMENT).			
GN	NRG1.			
ID	09ESAB	PRELIMINARY;	PR;	111 AA.
ID	09ESAB			
DT	01-MAR-2001 (TEMBREL. 16, Created)			
DT	01-MAR-2001 (TEMBREL. 16, Last sequence update)			
DT	01-DEC-2001 (TEMBREL. 19, Last annotation update)			
DE	SMDF NEUREGULIN BETA 2 (FRAGMENT).			
GN	NRG1.			
OS	Rattus norvegicus (Rat).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OT	01-MAR-2001 (TEMBREL. 16, Last sequence update)			
OT	01-MAR-2001 (TEMBREL. 16, Last sequence update)			
DE	SMDF NEUREGULIN BETA 2 (FRAGMENT).			
GN	NRG1.			
ID	09ESAB	PRELIMINARY;	PR;	111 AA.
ID	09ESAB			
DT	01-MAR-2001 (TEMBREL. 16, Created)			
DT	01-MAR-2001 (TEMBREL. 16, Last sequence update)			
DT	01-DEC-2001 (TEMBREL. 19, Last annotation update)			
DE	SMDF NEUREGULIN BETA 2 (FRAGMENT).			
GN	NRG1.			
OS	Rattus norvegicus (Rat).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OT	01-MAR-2001 (TEMBREL. 16, Last sequence update)			
OT	01-MAR-2001 (TEMBREL. 16, Last sequence update)			
DE	SMDF NEUREGULIN BETA 2 (FRAGMENT).			
GN	NRG1.			
ID	09ESAB	PRELIMINARY;	PR;	111 AA.
ID	09ESAB			
DT	01-MAR-2001 (TEMBREL. 16, Created)			
DT	01-MAR-2001 (TEMBREL. 16, Last sequence update)			
DT	01-DEC-2001 (TEMBREL. 19, Last annotation update)			
DE	SMDF NEUREGULIN BETA 2 (FRAGMENT).			
GN	NRG1.			
OS	Rattus norvegicus (Rat).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OT	01-MAR-2001 (TEMBREL. 16, Last sequence update)			
OT	01-MAR-2001 (TEMBREL. 16, Last sequence update)			
DE	SMDF NEUREGULIN BETA 2 (FRAGMENT).			
GN	NRG1.			
ID	09ESAB	PRELIMINARY;	PR;	111 AA.
ID	09ESAB			
DT	01-MAR-2001 (TEMBREL. 16, Created)			
DT	01-MAR-2001 (TEMBREL. 16, Last sequence update)			
DT	01-DEC-2001 (TEMBREL. 19, Last annotation update)			
DE	SMDF NEUREGULIN BETA 2 (FRAGMENT).			
GN	NRG1.			
OS	Rattus norvegicus (Rat).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OT	01-MAR-2001 (TEMBREL. 16, Last sequence update)			
OT	01-MAR-2001 (TEMBREL. 16, Last sequence update)			
DE	SMDF NEUREGULIN BETA 2 (FRAGMENT).			
GN	NRG1.			
ID	09ESAB	PRELIMINARY;	PR;	111 AA.
ID	09ESAB			
DT	01-MAR-2001 (TEMBREL. 16, Created)			
DT	01-MAR-2001 (TEMBREL. 16, Last sequence update)			
DT	01-DEC-2001 (TEMBREL. 19, Last annotation update)			
DE	SMDF NEUREGULIN BETA 2 (FRAGMENT).			
GN	NRG1.			
OS	Rattus norvegicus (Rat).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OT	01-MAR-2001 (TEMBREL. 16, Last sequence update)			
OT	01-MAR-2001 (TEMBREL. 16, Last sequence update)			
DE	SMDF NEUREGULIN BETA 2 (FRAGMENT).			
GN	NRG1.			
ID	09ESAB	PRELIMINARY;	PR;	111 AA.
ID	09ESAB			</

Query Match	37.0%;	Score 102.5;	DB 11;	Length 111;
Best Local Similarity	31.2%;	Pred. No. 8.1e-07;		
Matches 15; Conservative 13; Mismatches 19; Indels 1; Gaps 1;				



GN NRGI.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY;  
 RA Carroll S.L., Stonecypher M.S., Anderson K.D., Pearson R.J. Jr.,  
 RT Frohner P.W.,  
 \*Structural and Functional Diversity of Glial Growth Factor Isoforms  
 RT Expressed in Regenerating Peripheral Nerve and Associated Neurons.\*;  
 RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF194996; AAC28450.1; -.  
 DR HSSP; Q12784; IHRF.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR003598; IG\_c2.  
 DR InterPro; IPR003600; IG\_1like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; IG; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR SMART; SM00001; EGF\_1like; 1.  
 DR SMART; SM00409; EGF\_1like; 1.  
 DR SMART; SM00408; IGC2; 1.  
 DR SMART; SM00410; IG\_1like; 1.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 KW Immunoglobulin domain.  
 FT NON\_TER  
 SQ SEQUENCE 323 AA; 35358 MW; C7DF153A939A80C8 CRC64;

Query Match 37.0%; Score 102.5; DB 11; Length 323;  
 Best Local Similarity 31.2%; Pred. No. 2.5e-06;  
 Matches 15; Conservative 13; Mismatches 19; Indels 1; Gaps 1;

OY 1 HFKPCRDKLAYCLNDGECFYETLTGSHK-CRCCKEGYGVRCDOFL 47  
 DB 260 HLKCAEKERTFCVNGECFTVKDLSNPRLCKCPNEFTGDRCONYV 307

RESULT 11  
 Q9ESAI PRELIMINARY; PRT; 342 AA.  
 AC Q9ESAI;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update).  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE GLIAL GROWTH FACTOR GGF BETA 4 (FRAGMENT).  
 GN NRGI.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY;  
 RA Carroll S.L., Stonecypher M.S., Anderson K.D., Pearson R.J. Jr.,  
 RT Frohner P.W.,  
 \*Structural and Functional Diversity of Glial Growth Factor Isoforms  
 RT Expressed in Regenerating Peripheral Nerve and Associated Neurons.\*;  
 RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF194997; AAC28451.1; -.  
 DR HSSP; Q12784; IHRF.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR003598; IG\_c2.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; IG; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR SMART; SM00001; EGF\_1like; 1.  
 DR SMART; SM00409; IG; 1.

DR SMART; SM00408; IGC2; 1.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 KW Immunoglobulin domain.  
 FT NON\_TER  
 SQ SEQUENCE 342 AA; 37836 MW; 8BE3FC836553124 CRC64;

Query Match 37.0%; Score 102.5; DB 11; Length 342;  
 Best Local Similarity 31.2%; Pred. No. 2.7e-06;  
 Matches 15; Conservative 13; Mismatches 19; Indels 1; Gaps 1;

OY 1 HFKPCRDKLAYCLNDGECFYETLTGSHK-CRCCKEGYGVRCDOFL 47  
 DB 260 HLKCAEKERTFCVNGECFTVKDLSNPRLCKCPNEFTGDRCONYV 307

RESULT 12  
 Q9ESB1 PRELIMINARY; PRT; 700 AA.  
 AC Q9ESB1;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE SMOF NEUREGULIN BETA 1A.  
 GN NRGI.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY;  
 RA Carroll S.L., Anderson K.D., Frohner P.W.,  
 RT \*Structural and Functional Diversity of SMOF Neuregulin Splice  
 RT Variants Expressed in the Adult Rat Nervous System.\*;  
 RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF194438; AAC28427.1; -.  
 DR HSSP; Q12784; IHRF.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR002154; Neuregulin.  
 DR InterPro; IPR002114; PTS\_HPr\_ser.  
 DR Pfam; PF00008; EGF; 1.  
 DR Pfam; PF02158; Neuregulin; 1.  
 DR PRINTS; PR01089; NEUREGULIN.  
 DR SMART; SM00181; EGF; 1.  
 DR SMART; SM00001; EGF\_1like; 1.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE; PS00589; PTS\_HPr\_ser; UNKNOWN\_1.  
 SQ SEQUENCE 700 AA; 76386 MW; 2F811B17ECC49DA CRC64;

Query Match 37.0%; Score 102.5; DB 11; Length 700;  
 Best Local Similarity 31.2%; Pred. No. 5.7e-06;  
 Matches 15; Conservative 13; Mismatches 19; Indels 1; Gaps 1;

OY 1 HFKPCRDKLAYCLNDGECFYETLTGSHK-CRCCKEGYGVRCDOFL 47  
 DB 234 HLKCAEKERTFCVNGECFTVKDLSNPRLCKCPNEFTGDRCONYV 281

RESULT 13  
 Q9ESAS PRELIMINARY; PRT; 782 AA.  
 AC Q9ESAS;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE GLIAL GROWTH FACTOR BETA 1A (FRAGMENT).  
 GN NRGI.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=SPINAL CORD/BRAIN STEM;  
 RA Carroll S.L., Stonecypher M.S., Anderson K.D., Pearson R.J. Jr.,  
 RT Frohbert P.W.;  
 "Structural and Functional Diversity of Glial Growth Factor Isoforms  
 RT Expressed in Regenerating Peripheral Nerve and Associated Neurons.";  
 RT Submitted (Oct-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF194993; AAG28433.1;  
 DR HSSP: 012784; IHRE.  
 DR InterPro: IPR000561; EGF\_1like.  
 DR InterPro: IPR003599; Iq.  
 DR InterPro: IPR003598; Iq\_c2.  
 DR InterPro: IPR003006; Iq\_MHC.  
 DR InterPro: IPR002154; Neuregulin.  
 DR InterPro: IPR002114; Pts\_Hpr\_ser.  
 DR Pfam: PF00047; Iq; 1.  
 DR Pfam: PF02158; Neuregulin; 1.  
 DR PRINTS: PRO1089; NEUREGULIN.  
 DR SMART: SM00181; EGF; 1.  
 DR SMART: SM00001; EGF\_1like; 1.  
 DR SMART: SM00409; Iq; 1.  
 DR SMART: SM00408; IGC2; 1.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE: PS00589; Pts\_Hpr\_ser; UNKNOWN\_1.  
 DR KEGG: K01001; Immunoglobulin domain.  
 DR NON\_TER 1  
 SQ SEQUENCE 782 AA; 86036 MW; F6174A68F4E27BDE CRC64;

Query Match 37.0%; Score 102.5; DB 11; Length 782;  
 Best Local Similarity 31.2%; Pred. No. 6.5e-06;  
 Matches 15; Conservative 13; Mismatches 19; Indels 1; Gaps 1;

OY 1 HEKPCRDKLAVCLNDGCEVFIEITLGSKH-CRCCKEGYGVRCDFL 47  
 Db 316 HLIKCAKEKTECVANGGCEFTVKDLSNPSRYLCKPCNFTGDRCONYV 363

RESULT 14  
 O91M20 PRELIMINARY; PRT; 89 AA.  
 AC O91M20;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE LSDV016 EGF-LIKE GROWTH FACTOR.  
 GN LSDV016.  
 OS Lumpy skin disease virus.  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Capripoxvirus.  
 OX NCBI\_TaxID=59509;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NEERHLING 2490;  
 RX MEDLINE=21329495; PubMed=11435593;  
 RA Tulman E.R., Afonso C.L., Lu Z., Zsak L., Kutish G.F., Rock D.L.;  
 RT "Genome of lumpy skin disease virus.";  
 RL J. Virol. 75:7122-7130(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NEERHLING 2490;  
 RA Tulman E.R., Afonso C.L., Lu Z., Zsak L., Kutish G.F., Rock D.L.;  
 RT Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF325528; AAK84977.1;  
 SQ SEQUENCE 89 AA; 10646 MW; ID5F3FD7D06174E0 CRC64;

Query Match 35.7%; Score 99; DB 12; Length 89;  
 Best Local Similarity 43.9%; Pred. No. 2e-06;  
 Matches 18; Conservative 6; Mismatches 13; Indels 4; Gaps 1;

OY 7 DKDLAYCLNDGCEVFIEITLGSKH----HRCCKEGYGVRC 43  
 Db 40 DKSLNFCNLGTCYKTFITLISYKPKPLMFCRCCKLGYGVRC 80

RESULT 15  
 ID 001768 PRELIMINARY; PRT; 2192 AA.  
 AC 001768;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE HYPOTHETICAL 242.7 KDA PROTEIN.  
 GN 12183.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Du Z., Le T.T.;  
 RT "The sequence of C. elegans cosmid 721E3.";  
 RL Submitted (May-1997) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Waterston R.;  
 RT "Direct Submission.";  
 RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF003133; AAB54138.2;  
 DR HSSP: Q07954; ICR8.  
 DR InterPro: IPR000561; EGF\_1like.  
 DR InterPro: IPR002049; Laminin\_EGF.  
 DR InterPro: IPR000033; LDL\_receptor\_rep.  
 DR InterPro: IPR002172; LDL\_receptor\_A.  
 DR Pfam: PF00008; EGF; 8.  
 DR Pfam: PF00057; LDL\_receptor\_a; 18.  
 DR Pfam: PF00058; LDL\_receptor\_b; 8.  
 DR PRINTS: PRO0011; EGF\_LAMININ.  
 DR PRINTS: PRO0261; LDLRECEPTOR.  
 DR SMART: SM00181; EGF; 23.  
 DR SMART: SM00192; LDLa; 18.  
 DR SMART: SM00135; IY; 7.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_9.  
 DR PROSITE: PS01186; EGF\_2; 7.  
 DR PROSITE: PS01209; LDLRA\_1; 10.  
 DR PROSITE: PS00068; LDLRA\_2; 12.  
 KW EGF-like domain; Glycoprotein; Hypothetical protein.  
 SQ SEQUENCE 2192 AA; 242666 MW; F4762A5EBCA45BDA CRC64;

Query Match 33.4%; Score 92.5; DB 5; Length 2192;  
 Best Local Similarity 42.9%; Pred. No. 0.00049;  
 Matches 18; Conservative 8; Mismatches 7; Indels 9; Gaps 3;  
 OY 5 CRDKLAYCLNDGCEVFIEITLGSKHRCCKEGYGVRCDFL 45  
 Db 1918 CDD----YCTNNSKC-----TITNGTHFECDCRKGFKLNCQ 1951

Search completed: September 13, 2002, 08:55:27  
 Job time: 354 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 08:49:53 ; Search time 10.36 Seconds

(without alignments)  
175.658 million cell updates/sec

Title: US-09-877-665-4

Perfect score: 277  
1 HFRPCRKDLAYCLNDGECE.....SHKRCCKEGYGVRCDFL 47

Sequence: 1 HFRPCRKDLAYCLNDGECE.....SHKRCCKEGYGVRCDFL 47

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	277	100.0	713 1	NRG3_MOUSE
2	277	100.0	720 1	NRG3_HUMAN
3	126.5	45.7	677 1	NRG1_XENLA
4	113.5	41.0	639 1	NRG1_HUMAN
5	110	39.7	115 1	NRG4_MOUSE
6	104.5	37.7	296 1	SMDE_HUMAN
7	104	37.5	756 1	NRG2_MOUSE
8	102.5	37.0	662 1	NRG1_RAT
9	98.5	35.6	602 1	NRG1_CHICK
10	90	32.5	850 1	NRG2_HUMAN
11	90	32.5	868 1	NRG2_RAT
12	89.5	32.3	169 1	EREG_HUMAN
13	85	30.7	80 1	GRFA_SFVKA
14	83	30.0	2531 1	NRG1_MOUSE
15	82.5	29.8	177 1	ETC_MOUSE
16	82.5	29.8	178 1	ETC_BOVIN
17	82	29.6	85 1	GRFA_MYXVL
18	82	29.6	230 1	SPIT_DROME
19	80.5	29.1	125 1	NRG1_MOUSE
20	80.5	29.1	178 1	ETC_HUMAN
21	79.5	28.7	140 1	GRFA_VACCV
22	79.5	28.7	142 1	GRFA_VACCC
23	79.5	28.7	159 1	GRFA_MOUSE
24	79.5	28.7	159 1	GRFA_MOUSE
25	79	28.5	2531 1	NRG1_RAT
26	78	28.2	1207 1	EGF_HUMAN
27	77.5	28.0	714 1	DLIL_RAT
28	77.5	28.0	722 1	DLIL_MOUSE
29	76	27.4	484 1	EGF_PIG
30	76	27.4	1217 1	EGF_MOUSE
31	76	27.1	2139 1	CRB_DROME
32	75	27.1	611 1	LEM2_CANFA
33	74.5	26.9	603 1	FAI2_CAVPO

34	74.5	26.9	723 1	DLIL_HUMAN	000548 homo sapien
35	74.5	26.9	1429 1	L112_CAEL	P14585 caenorhabd
36	74	26.7	551 1	LEM2_RABIT	P27113 oryctolagus
37	74	26.7	610 1	LEM2_HUMAN	P16581 homo sapien
38	74	26.7	2871 1	FBNI_BOVIN	P98133 bos taurus
39	74	26.7	2871 1	FBNI_HUMAN	P33555 homo sapien
40	74	26.7	2871 1	FBNI_MOUSE	Q61554 mus musculu
41	74	26.7	2871 1	FBNI_PIG	Q9tv36 sus scrofa
42	73	26.4	294 1	GRK_DROME	P42287 drosophila
43	73	26.4	409 1	MGK_PIG	P79385 sus scrofa
44	73	26.4	1964 1	NTC4_MOUSE	P31695 mus musculu
45	73	26.4	2524 1	NOTC_XENLA	P21783 xenopus lae

## ALIGNMENTS

RESULT 1  
ID NRG3\_MOUSE STANDARD; PRT; 713 AA.  
AC 035181;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Pro-neuregulin-3 precursor (Pro-NRG3) [Contains: Neuregulin-3 (NRG-3)].  
GN NRG3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA MEDLINE=97420720; PubMed=9275162;  
RA Zhang D., Sliwkowski M.X., Mark M., Frantz G., Akita R., Sun Y., Hillan K., Crowley C., Brush J., Godowski P.J.;  
RT "Neuregulin-3 (NRG3): a novel neural tissue-enriched protein that binds and activates ErbB4".  
RL Proc. Natl. Acad. Sci. U.S.A. 94:9562-9567(1997).  
-1- FUNCTION: DIRECT LIGAND FOR THE ERBB4 TYROSINE KINASE RECEPTOR. BINDING RESULTS IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND ACTIVATION OF THE RECEPTOR. DOES NOT BIND TO THE EGF RECEPTOR, ERBB2 OR ERBB3 RECEPTORS.  
-1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).  
-1- TISSUE SPECIFICITY: EXPRESSED IN SYMPATHETIC, MOTOR, AND SENSORY NEURONS.  
-1- DEVELOPMENTAL STAGE: DETECTED AS EARLY AS E11. IN E13 EMBRYOS, DETECTED MAINLY IN THE NERVOUS SYSTEM. IN E16 EMBRYOS, DETECTED IN THE BRAIN, SPINAL CORD, TRIGEMINAL, VESTIBULAR-COCHLEAR, AND SPINAL GANGLIA. IN ADULTS, EXPRESSED IN SPINAL CORD, AND NUMEROUS BRAIN REGIONS.  
-1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN DIMERIZATION (BY SIMILARITY).  
-1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE DOMAIN (BY SIMILARITY).  
-1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR FORM (BY SIMILARITY).  
-1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY SIMILARITY).  
-1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
-1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.  
-----  
This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: AF010130; AAB70914.1; --  
 CC MGD: MGI:1097165; Nf93.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR002154; Neuregulin.  
 DR Pfam: PF00008; EGF; 1.  
 DR Pfam: PF02158; Neuregulin; 1.  
 DR SMART: SM00181; EGF; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 KM Growth factor; EGF-like domain; Transmembrane; Multigene family.  
 FT CHAIN 1 713 PRO-NEUREGULIN-3, MEMBRANE-BOUND FORM.  
 FT CHAIN 1 362 NEUREGULIN-3.  
 FT DOMAIN 1 362 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 363 383 INTERNAL SIGNAL SEQUENCE (POTENTIAL).  
 FT DOMAIN 384 713 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 105 287 SER/THR-RICH.  
 FT DOMAIN 288 331 EGF-LIKE.  
 FT DOMAIN 13 21 POLY-ALA.  
 FT DOMAIN 26 34 POLY-ALA.  
 FT DOMAIN 127 135 POLY-THR.  
 FT DOMAIN 250 253 POLY-ALA.  
 FT DOMAIN 254 263 POLY-SER.  
 FT DOMAIN 264 267 POLY-THR.  
 FT DISULFID 292 306 BY SIMILARITY.  
 FT DISULFID 300 319 BY SIMILARITY.  
 FT DISULFID 321 330 BY SIMILARITY.  
 SQ SEQUENCE 713 AA; 77369 MW; 9E7D1D5E7FC8DC90 CRC64;

Query Match 100.0%; Score 277; DB 1; Length 713;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-26;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFKPCRDKLAVCLNDGCEFYIEITLTGSHKRCRKEGYGVRCDFL 47  
 DB 288 HFKPCRDKLAVCLNDGCEFYIEITLTGSHKRCRKEGYGVRCDFL 334

RESULT 2  
 NRG3\_HUMAN STANDARD; PRT; 720 AA.  
 AC P56975;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Pro-neuregulin-3 precursor (Pro-NRG3) [Contains: Neuregulin-3 (NRG-3)].  
 GN NRG3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal brain;  
 RX MEDLINE=97420720; PubMed=9275162;  
 RA Zhang D., Sliwkowski M.X., Mark M., Frantz G., Akita R., Sun Y.,  
 RA Hillan K., Crowley C., Brush J., Godowski P.J.;  
 RA "Neuregulin-3 (NRG3): a novel neural tissue-enriched protein that  
 binds and activates ErbB4.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 94:9562-9567(1997).  
 RL -1- FUNCTION: DIRECT LIGAND FOR THE ERBB4 TYROSINE KINASE RECEPTOR.  
 CC BINDING RESULTS IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND  
 CC ACTIVATION OF THE RECEPTOR. DOES NOT BIND TO THE EGF RECEPTOR.  
 CC ERBB2 OR ERBB3 RECEPTORS.  
 CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS  
 CC A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE  
 CC MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN ACTIVE REGIONS OF THE BRAIN  
 CC WITH THE EXCEPTION OF CORPUS CALLOSUM. EXPRESSED AT LOWER LEVEL IN

TESTIS. NOT DETECTED IN HEART, PLACENTA, LUNG, LIVER, SKELETAL  
 CC MUSCLE, KIDNEY, PANCREAS, SPLEEN, THYMUS, PROSTATE, OVARY, SMALL  
 CC INTESTINE, COLON AND PERIPHERAL BLOOD LEUKOCYTES.  
 CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION  
 CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE  
 CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN  
 CC DIMERIZATION (BY SIMILARITY).  
 CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE  
 CC DOMAIN (BY SIMILARITY).  
 CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE  
 CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR  
 CC FORM (BY SIMILARITY).  
 CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.  
 DR MIM: 605533; --  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR002154; Neuregulin.  
 DR Pfam: PF02158; Neuregulin; 1.  
 DR SMART: SM00181; EGF; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 KM Growth factor; EGF-like domain; Transmembrane; Multigene family.  
 FT CHAIN 1 720 PRO-NEUREGULIN-3, MEMBRANE-BOUND FORM.  
 FT CHAIN 1 359 NEUREGULIN-3.  
 FT DOMAIN 1 360 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 361 381 INTERNAL SIGNAL SEQUENCE (POTENTIAL).  
 FT DOMAIN 382 720 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 105 285 SER/THR-RICH.  
 FT DOMAIN 286 329 EGF-LIKE.  
 FT DOMAIN 5 8 POLY-ALA.  
 FT DOMAIN 13 21 POLY-ALA.  
 FT DOMAIN 26 34 POLY-ALA.  
 FT DOMAIN 127 135 POLY-THR.  
 FT DOMAIN 252 260 POLY-SER.  
 FT DOMAIN 265 265 POLY-THR.  
 FT DISULFID 290 304 BY SIMILARITY.  
 FT DISULFID 298 317 BY SIMILARITY.  
 FT DISULFID 319 328 BY SIMILARITY.  
 SQ SEQUENCE 720 AA; 77900 MW; AAD6F10DB95A693 CRC64;

Query Match 100.0%; Score 277; DB 1; Length 720;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-26;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFKPCRDKLAVCLNDGCEFYIEITLTGSHKRCRKEGYGVRCDFL 47  
 DB 286 HFKPCRDKLAVCLNDGCEFYIEITLTGSHKRCRKEGYGVRCDFL 332

RESULT 3  
 NRG1\_XENLA STANDARD; PRT; 677 AA.  
 AC O93383; Q9W6N0;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Pro-neuregulin-1 precursor (Pro-NRG1) [Contains: Neuregulin-1].  
 GN NRG1.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;  
 OC Xenopodidae; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA1), AND ALTERNATIVE SPLICING.  
 RX MEDLINE=98352126; PubMed=9685585;  
 RA Yang J.F., Zhou H., Pun S., Ip N.Y., Peng H.B., Tsai K.W.K.;  
 RA "Cloning of cDNAs encoding xenopus neuregulin: expression in myotomal  
 muscle during embryo development.";  
 RT Brain Res. Mol. Brain Res. 58:59-73(1998).  
 RL Brain Res. Mol. Brain Res. 58:59-73(1998).



RC TISSUE-BRAIN: Pubmed-8096067; Chen M.S., Berningham-McDonogh O.,  
 RA MEDLINE-93205115; Goodert A.D.J., Danehy F., Mismil D., Sudhalter J.,  
 RA Marchionni M.A., Hendricks M., Danehy F., Mismil D., Sudhalter J.,  
 RA Kik C., Hendricks M., Danehy F., Mismil D., Sudhalter J.,  
 RA Koyashiki K., Wroblewski D., Lynch C., Balasare M., Hiles I.,  
 RA Wasser J.B., Hsuan J.J., Rott N.F., Otsu M., McBurney R.N.,  
 RA Waterfield M.D., Stroobant P., Gwynne D.;  
 RT "Glial growth factors are alternatively spliced erb2 ligands  
 RT expressed in the nervous system.";  
 RL Nature 362:312-318(1993).  
 RN [5]  
 RP SEQUENCE FROM N.A. OF GAMMA-HEREGULIN FUSION PROTEIN.  
 RC TISSUE-Breast cancer; Pubmed-9333014;  
 RA MEDLINE-97472144; Schaefer G., Fitzpatrick V.D., Sliwowski M.X.;  
 RA Schaefer G., Fitzpatrick V.D., Sliwowski M.X.;  
 RT "Gamma-hergulin: a novel heregulin isoform that is an autocrine  
 RT growth factor for the human breast cancer cell line, MDA-MB-175.";  
 RN Oncogene 15:1385-1394(1997).  
 RL [6]  
 RP SEQUENCE OF 1-210 FROM N.A.  
 RA Schoumacher F., Herzer S., Flury N., Kueng W., Mueller H.,  
 RA Spenner U.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE OF 19-27.  
 RA MEDLINE-93366731; Pubmed-7689552;  
 RA Culouscou J.M., Plowman G.D., Carlton G.W., Green J.M., Shoyab M.;  
 RT "Characterization of a breast cancer cell differentiation factor that  
 RT specifically activates the HER4/p180erbB4 receptor.";  
 RN J. Biol. Chem. 268:18407-18410(1993).  
 RL [8]  
 RP CHROMOSOMAL TRANSLOCATION.  
 RA MEDLINE-99455251; Pubmed-10523851;  
 RA Wang X.-Z., Jolicoeur E.M., Conte N., Chaffanet M., Zhang Y.,  
 RA Morziconacci M.-J., Fehner H., Birnbaum D., Pedusque M.-J., Ron D.;  
 RT "Gamma-hergulin is the product of a chromosomal translocation fusing  
 RT the DC4 and HGL/NR1 genes in the MDA-MB-175 breast cancer cell  
 RT line.";  
 RL Oncogene 18:5718-5721(1999).  
 RN [9]  
 RP CHROMOSOMAL TRANSLOCATION.  
 RA MEDLINE-20065180; Pubmed-10597312;  
 RA Liu X., Baker E., Eyre H.J., Sutherland G.R., Zhou M.;  
 RT "Gamma-hergulin: a fusion gene of DOC-4 and neueregulin-1 derived from  
 RT a chromosome translocation.";  
 RL Oncogene 18:7110-7114(1999).  
 RN [10]  
 RP STRUCTURE BY NMR OF 175-241 (ISOFORM ALPHA).  
 RA MEDLINE-94341264; Pubmed-8062828;  
 RA Nagata K., Kohda D., Hatanaka H., Ichikawa S., Matsuda S.,  
 RA Yamamoto T., Suzuki A., Inagaki F.;  
 RT "Solution structure of the epidermal growth factor-like domain of  
 RT heregulin-alpha, a ligand for p180erbB-4.";  
 RL EMBO J. 13:3517-3523(1994).  
 CC -1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE  
 CC RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,  
 CC RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND  
 CC ACTIVATION OF THE ERBB RECEPTORS. THE MULTIPLE ISOFORMS PERFORM  
 CC DIVERSE FUNCTIONS SUCH AS INDUCING GROWTH AND DIFFERENTIATION OF  
 CC EPITHELIAL, GLIAL, NEURONAL, AND SKELETAL MUSCLE CELLS; INDUCING  
 CC EXPRESSION OF ACETYLCHOLINE RECEPTOR IN SYNAPTIC VESICLES DURING  
 CC THE FORMATION OF THE NEUROMUSCULAR JUNCTION; STIMULATING  
 CC LOBULOVULAR BUDDING AND MILK PRODUCTION IN THE MAMMARY GLAND  
 CC AND INDUCING DIFFERENTIATION OF MAMMARY TUMOR CELLS; STIMULATING  
 CC SCHWANN CELL PROLIFERATION; IMPLICATION IN THE DEVELOPMENT OF THE  
 CC MYOCARDIUM SUCH AS TRABECULATION OF THE DEVELOPING HEART.  
 CC -1- SUBUNIT: THE CYTOPLASMIC DOMAIN INTERACTS WITH THE LIM DOMAIN  
 CC REGION OF LIMK1 (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS  
 CC A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE  
 CC MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE. THE SECRETED  
 CC ISOFORM GGF2 HAS A SIGNAL PEPTIDE. THE ISOFORM BETA3 MAY BE  
 CC NUCLEAR.

CC -1- ALTERNATIVE PRODUCTS: AT LEAST 10 ISOFORMS; ALPHA (SHOWN HERE);  
 CC ALPHA1A, ALPHA2B, ALPHA3, BETA1, BETA2, BETA3/GGFEB1,  
 CC GGF2/GGFEB2 AND SMDF (AC 015491); ARE PRODUCED BY ALTERNATIVE  
 CC SPLICING. THEY HAVE BEEN CLASSIFIED AS TYPE I NRGS (VARIANTS WITH  
 CC AN IG DOMAIN AND A GLYCOSYLATION DOMAIN; ALPHA AND BETA), TYPE II  
 CC NRGS (VARIANTS WITH AN IG DOMAIN BUT NO GLYCOSYLATION DOMAIN;  
 CC GGF2) AND TYPE III NRGS (VARIANTS WITH A CYS-RICH DOMAIN; SMDF).  
 CC ALL THESE ISOFORMS PERFORM DISTINCT TISSUE-SPECIFIC FUNCTIONS.  
 CC -1- TISSUE SPECIFICITY: TYPE I ISOFORMS ARE THE PREDOMINANT FORMS  
 CC EXPRESSED IN THE ENDOCARDIUM. ISOFORM ALPHA IS EXPRESSED IN  
 CC BREAST, OVARY, TESTIS, PROSTATE, HEART, SKELETAL MUSCLE, LUNG,  
 CC PLACENTA LIVER, KIDNEY, SALIVARY GLAND, SMALL INTESTINE AND BRAIN,  
 CC BUT NOT IN UTERUS, STOMACH, PANCREAS, AND SPLEEN. ALPHA2 IS THE  
 CC PREDOMINANT FORM IN MESENCHYMAL CELLS AND IN NONNEURONAL ORGANS,  
 CC WHEREAS BETA1 IS THE MAJOR NEURONAL FORM. BETA3 IS EXPRESSED IN  
 CC SPINAL CORD AND BRAIN. GGF2 IS THE MAJOR FORM IN SKELETAL MUSCLE  
 CC CELLS; IN THE NERVOUS SYSTEM IT IS EXPRESSED IN SPINAL CORD AND  
 CC BRAIN, ALSO DETECTED IN ADULT HEART, PLACENTA, LUNG, LIVER,  
 CC KIDNEY, AND PANCREAS.  
 CC -1- DEVELOPMENTAL STAGE: DETECTABLE AT EARLY EMBRYONIC AGES.  
 CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION  
 CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE  
 CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN  
 CC DIMERIZATION (BY SIMILARITY).  
 CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE  
 CC DOMAIN.  
 CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE  
 CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR  
 CC FORM.  
 CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY  
 CC SIMILARITY).  
 CC -1- DISEASE: INVOLVED IN A RARE T(8;11) CHROMOSOMAL TRANSLOCATION THAT  
 CC FUSES THE 5' END OF OD24 TO NRG1 (ISOFORM BETA3). THE PRODUCT OF  
 CC THIS TRANSLOCATION HAS FIRST BE THOUGHT TO BE AN ALTERNATIVE  
 CC SPLICED ISOFORM, CALLED GAMMA-HEREGULIN. GAMMA-HEREGULIN IS A  
 CC SOLUBLE ACTIVATING LIGAND FOR THE ERBB2-ERBB3 RECEPTOR COMPLEX AND  
 CC ACTS AS AN AUTOCRINE GROWTH FACTOR IN A SPECIFIC BREAST CANCER  
 CC CELL LINE (MDA-MB-175). NOT DETECTED IN BREAST CARCINOMA SAMPLES,  
 CC INCLUDING DUCTAL, LOBULAR, MEDULLARY, AND MUCINOUS HISTOLOGICAL  
 CC TYPES. NEITHER IN OTHER BREAST CANCER CELL LINES.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: M94165; AAA58638.1; -  
 CC EMBL: M94166; AAA58639.1; -  
 CC EMBL: M94167; AAA58640.1; -  
 CC EMBL: M94168; AAA58641.1; -  
 CC EMBL: L12261; AAB59358.1; -  
 CC EMBL: U02325; AAA19850.1; -  
 CC EMBL: U02326; AAA19951.1; -  
 CC EMBL: U02327; AAA19952.1; -  
 CC EMBL: U02328; AAA19953.1; -  
 CC EMBL: U02329; AAA19954.1; -  
 CC EMBL: U02330; AAA19955.1; -  
 CC EMBL: L12260; AAB59622.1; -  
 CC EMBL: AF026146; AAD01795.1; -  
 CC EMBL: AF009227; AAC51756.1; ALT\_INIT.  
 CC PDB: 1HRE; 15-OCT-94.  
 CC PDB: 1HRE; 15-OCT-94.  
 CC MIM: 142445; -  
 CC InterPro: IPR000561; EGF-like.  
 CC InterPro: IPR003006; Ig\_MHC.  
 CC InterPro: IPR002154; Neuregulin.  
 CC Pfam: PF00047; Ig; 1.

DR Pfam: PF02158; Neuregulin-1.  
 DR PRINTS: PRO1089; NEUREGULIN.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein;  
 KW Transmembrane; Multigene family; Polymorphism; 3D-structure;  
 KW Alternative splicing; Chromosomal translocation.

Query Match 41.0%; Score 113.5; DB 1; Length 639;  
 Best Local Similarity 34.8%; Pred. No. 2,3e-06;  
 Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

OY 1 HFKCRDLDLAVCLNDGCEYIEITLGSKH-CCKCKEYQGVRCQ 45  
 Db 177 HLVCAKERKTECVNGGCEVFKDLSNPSRYLCKQPGFTGRCETE 222

RESULT 5  
 NRGA\_MOUSE STANDARD; PRT; 115 AA.  
 ID 09MTX4;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Pro-neuregulin-4, short isoform (Pro-NRG4) [Contains: Neuregulin-4  
 (NRG-4)].  
 GN NRG4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Liver;  
 RX MEDLINE=99276098; PubMed=10348342;  
 RA Harari D., Tzahar E., Romano J., Shelly M., Pierce J.H., Andrews G.C.,  
 RT Yarden Y.;  
 RA "Neuregulin-4: a novel growth factor that acts through the ErbB-4  
 receptor tyrosine kinase.";  
 RL Oncogene 18:2681-2689(1999).  
 CC -1- FUNCTION: LOW AFFINITY LIGAND FOR THE ERBB4 TYROSINE KINASE  
 RECEPTOR. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,  
 RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND  
 ACTIVATION OF THE ERBB RECEPTORS. DOES NOT BIND TO THE ERBB1,  
 ERBB2 AND ERBB3 RECEPTORS.  
 CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS  
 A PROTOPOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE  
 MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).  
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS MAY BE PRODUCED BY  
 ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PANCREAS; WEAKLY EXPRESSED  
 IN MUSCLE.  
 CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION  
 OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE  
 PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN  
 DIMERIZATION (BY SIMILARITY).  
 CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE  
 DOMAIN (BY SIMILARITY).  
 CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE  
 EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR  
 FORM (BY SIMILARITY).  
 CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY  
 SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.  
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: AF083067; AAD21874.1; -  
 DR MGD; MGI:1933833; Nrg4.  
 DR InterPro: IPR000561; EGF\_1-like.  
 DR InterPro: IPR001336; EGF\_1.  
 DR Pfam: PF00008; EGF; 1.  
 DR PRINTS: PRO0009; EGF\_TF.  
 DR SMART: SM00181; EGF; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; FALSE\_NEG.  
 DR Growth factor; EGF-like domain; Glycoprotein; Transmembrane;  
 KW Multigene family; Alternative splicing.  
 KW CHAIN 1 115  
 FT CHAIN 1 61  
 FT DOMAIN 1 62  
 FT TRANSMEM 63 83  
 FT DOMAIN 84 115  
 FT DOMAIN 5 46  
 FT DISULFID 9 23  
 FT DISULFID 17 34  
 FT DISULFID 36 45  
 FT CARBOHYD 39 39  
 FT CARBOHYD 60 60  
 SQ SEQUENCE 115 AA; 12743 MW; 989A1E376F857B49 CRC64;

Query Match 39.7%; Score 110; DB 1; Length 115;  
 Best Local Similarity 42.2%; Pred. No. 1,2e-06;  
 Matches 19; Conservative 8; Mismatches 16; Indels 2; Gaps 1;

OY 1 HFKCRDLDLAVCLNDGCEYIEITLGSKHCKCKEYQGVRCQ 45  
 Db 5 HEQPCGPRHRSFCLNGICGYIPIPS--PPRCRIENVTGARCEE 47

RESULT 6  
 SMDP\_HUMAN STANDARD; PRT; 296 AA.  
 ID SMDP\_HUMAN  
 AC 015491;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Neuregulin-1, sensory and motor neuron-derived factor isoform.  
 GN NG1 OR HGL OR NDF OR HRGA OR GGF OR SMDP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain stem, and Cerebellum;  
 RX MEDLINE=95301541; PubMed=7782315;  
 RA Ho W.-H., Aramant M.P., Nuljens A., Phillips H.S., Osheroff P.L.;  
 RT "Sensory and motor neuron-derived factor. A novel heregulin variant  
 highly expressed in sensory and motor neurons.";  
 RL J. Biol. Chem. 270:14523-14532(1995).  
 CC -1- FUNCTION: THE ISOFORM SMDP MAY PLAY A ROLE IN MOTOR AND SENSORY  
 NEURON DEVELOPMENT.  
 CC -1- SUBCELLULAR LOCATION: SECRETED. MAY POSSESS AN INTERNAL UNCLEAVED  
 SIGNAL SEQUENCE.  
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 10 ISOFORMS OF NG1 ARE PRODUCED BY  
 ALTERNATIVE SPLICING. EXCEPT FOR SMDP THEY ARE IN ENTRY AC 002297.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN NERVOUS SYSTEM: SPINAL CORD MOTOR  
 NEURONS, DORSAL ROOT GANGLION NEURONS, AND BRAIN. PREDOMINANT  
 ISOFORM EXPRESSED IN SENSORY AND MOTOR NEURONS. NOT DETECTED IN  
 ADULT HEART, PLACENTA, LUNG, LIVER, SKELETAL MUSCLE, KIDNEY, AND  
 PANCREAS. NOT EXPRESSED IN FETAL LUNG, LIVER, AND KIDNEY.  
 CC -1- DEVELOPMENTAL STAGE: HIGHLY EXPRESSED IN DEVELOPING SPINAL MOTOR  
 NEURONS AND IN DEVELOPING CRANIAL NERVE NUCLEI. EXPRESSION IS  
 MAINTAINED ONLY IN BOTH ADULT MOTOR NEURONS AND DORSAL ROOT  
 GANGLION NEURONS.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: LA1827; AAC41764.1; -  
 DR HSSP: Q12784; 1HRE.  
 DR MIM: 142445; -  
 DR InterPro: IPR000551; EGF-like.  
 DR Pfam: PF00008; EGF\_1.  
 DR SMART: SM00181; EGF\_1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; FALSE\_NEG.  
 KW Growth factor; EGF-like domain; Transmembrane; Multigene family;  
 KM Alternative splicing.  
 FT TRANSMEM 76 100 INTERNAL SIGNAL SEQUENCE (POTENTIAL).  
 FT DOMAIN 58 91 CYS-RICH.  
 FT DOMAIN 211 232 SER/THR-RICH.  
 FT DISULFID 237 251 EGF-LIKE.  
 FT DISULFID 245 265 BY SIMILARITY.  
 FT DISULFID 267 276 BY SIMILARITY.  
 SQ SEQUENCE 296 AA; 31685 MW; 8D41743217F7EB02 CRC64;  
 OY 1 HRPCHDKDLAYCLNDGECFVETLTGSHKRCCKEGYGVRCDFL 47  
 DB 233 HLYKCAKEKTEFCVNGEGFMKDLNSRYLCKCPNETTGRCONYV 280  
 Query Match 37.7%; Score 104.5; DB 1; Length 296;  
 Best Local Similarity 31.2%; Pred. No. 1.3e-05;  
 Matches 15; Conservative 14; Mismatches 18; Indels 1; Gaps 1;  
 RESULT 7  
 NR02\_MOUSE STANDARD; PRF; 756 AA.  
 AC P56974;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Pro-neuregulin-2 precursor (Pro-NRG2) [Contains: Neuregulin-2 (NRG-2)  
 DE (Divergent of neuregulin 1) (DON-1)].  
 GN NRG2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS NRG2-5; NRG2-10 AND NRG2-16A).  
 RC STRAIN-C57BL/6; TISSUE-Brain.  
 RX MEDLINE=9731398; PubMed=9168115;  
 RA Carraway K.L., Ili, Weber J.L., Unger M.J., Ledesma J., Yu N.,  
 RA Gassmann M., Lai C.;  
 RT "Neuregulin-2, a new ligand of ErbB3/ErbB4-receptor tyrosine  
 RT kinases";  
 RT Nature 387:512-516(1997).  
 RL [2]  
 RP SEQUENCE OF 150-756 FROM N.A. (ISOFORMS DON-1M AND DON-1S).  
 RC TISSUE-Choroid plexus;  
 RX MEDLINE=97342638; PubMed=9193335;  
 RA Busfield S.J., Michnick D.A., Chikering T.W., Revett T.L., Ma J.,  
 RA Woolf E.A., Comrack C.A., Dussault B.J., Woolf J., Goodearl A.D.J.,  
 RA Gearing D.P.;  
 RT "Characterization of a neuregulin-related gene, Don-1, that is highly  
 RT expressed in restricted regions of the cerebellum and hippocampus";  
 RL Mol. Cell. Biol. 17:4007-4014(1997).  
 CC -1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE  
 CC RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORRECEPTORS,

CC RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND  
 CC ACTIVATION OF THE ERBB RECEPTOR. MAY ALSO PROMOTE THE  
 CC HETERO-DIMERIZATION WITH THE EGF RECEPTOR.  
 CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS  
 CC A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE  
 CC MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).  
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS: DON-1M, DON-1S/NRG2-5,  
 CC NRG2-10 AND NRG2-16A (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE  
 CC SPLICING.  
 CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN THE BRAIN, WITH LOWER  
 CC LEVELS IN THE LUNG. IN THE CEREBELLUM, FOUND IN GRANULE AND  
 CC PURKINJE CELLS.  
 CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION  
 CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE  
 CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN  
 CC DIMERIZATION (BY SIMILARITY).  
 CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE  
 CC DOMAIN (BY SIMILARITY).  
 CC -1- PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE  
 CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR  
 CC FORM (BY SIMILARITY).  
 CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.  
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.  
 CC MGD: MGI:1098246; Nrg2.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR003006; I9-MHC.  
 DR InterPro: IPR003598; I9-C2.  
 DR InterPro: IPR002154; Neuregulin.  
 DR Pfam: PF00047; I9; 1.  
 DR Pfam: PF02158; Neuregulin; 2.  
 DR SMART: SM00181; EGF\_1.  
 DR SMART: SM00408; IGC2; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 KW Growth factor; EGF-like domain; Immunoglobulin domain; glycoprotein;  
 KW Transmembrane; Multigene family; Alternative splicing.  
 KM Transmembrane; Multigene family; Alternative splicing.  
 FT PROPEP 1 19  
 FT CHAIN 20 756  
 FT CHAIN 20 314  
 FT CHAIN 20 315  
 FT DOMAIN 20 315  
 FT TRANSMEM 316 336  
 FT DOMAIN 337 756  
 FT DOMAIN 158 226  
 FT DOMAIN 238 248  
 FT DOMAIN 249 290  
 FT DOMAIN 627 633  
 FT DISULFID 165 219  
 FT DISULFID 253 267  
 FT DISULFID 261 278  
 FT DISULFID 280 289  
 FT CARBOHYD 55 55  
 FT CARBOHYD 186 186  
 FT CARBOHYD 254 254  
 FT CARBOHYD 296 296  
 FT CARBOHYD 280 280  
 FT VARSPLIC 281 756  
 FT VARSPLIC 282 330  
 FT VARSPLIC 331 756  
 FT VARSPLIC 282 307  
 FT VARSPLIC 307 756  
 SQ SEQUENCE 756 AA; 82213 MW; 51D85DC91BBE678E CRC64;  
 OY 1 HRPCHDKDLAYCLNDGECFVETLTGSHKRCCKEGYGVRCDF 46  
 Query Match 37.5%; Score 104; DB 1; Length 756;  
 Best Local Similarity 39.1%; Pred. No. 3.7e-05;  
 Matches 18; Conservative 8; Mismatches 18; Indels 2; Gaps 1;



Db 249 HARCNETAKSYCVNGVCYIEBT--NOLSKCPVGTGRCQOF 292

RESULT 8

NRG1\_RAT STANDARD: PRT: 662 AA.  
ID NRG1\_RAT  
P33322; P43323; P43324; P43325; P43326; P43327; P43328;  
16-OCT-2001 (Rel. 40, Created)  
16-OCT-2001 (Rel. 40, Last sequence update)  
16-OCT-2001 (Rel. 40, Last annotation update)  
Pro-neuregulin-1 precursor (Pro-NRG) [Contains: Neuregulin-1 (New differentiation factor) (Heregulin) (HRG) (Acetylcholine receptor inducing activity) (ARIA) (Sensory and motor neuron-derived factor) (Glia growth factor)].  
NRG1 OR NDF.  
Rattus norvegicus (Rat).  
Chordata; Craniata; Vertebrata; Euteleostomi; Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
NCBI\_TaxID=10116;  
[1]  
SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
TISSUE=Embryonic;  
MEDLINE=94158863; PubMed=7509448;  
Wen D., Suga S.V., Karunagaran D., Liu N., Cupples R.L., Luo Y., Janssen A.M., Ben-Baruch N., Trollinger D.B., Jacobsen V.L., Meng S.-Y., Lu H.S., Hu S., Chang D., Yang W., Yanagihara D., Koski R.A., Yarden Y.;  
"Structural and functional aspects of the multiplicity of Neu differentiation factors.";  
Mol. Cell. Biol. 14:1909-1919(1994).  
[2]  
SEQUENCE FROM N.A. (ISOFORM ALPHA2C/NDP44), AND PARTIAL SEQUENCE.  
TISSUE=Embryonic;  
MEDLINE=9225796; PubMed=1349853;  
Wen D., Peles E., Cupples R., Suga S.V., Bacus S.S., Luo Y., Trull G., Hu S., Silbiger S.M., Levy R.B., Koski R.A., Lu H.S., Yarden Y.;  
"Neu differentiation factor: a transmembrane glycoprotein containing an EGF domain and an immunoglobulin homology unit.";  
Cell 69:559-572(1992).  
[3]  
SEQUENCE OF 14-36.  
MEDLINE=92208945; PubMed=1348215;  
Peles E., Bacus S.S., Koski R.A., Lu H.S., Wen D., Ogdén S.G., Levy R.B., Yarden Y.;  
"Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein that induces differentiation of mammary tumor cells.";  
Cell 69:205-216(1992).  
[4]  
REGULATION OF PROCESSING (ISOFORM ALPHA2C/NDP44).  
MEDLINE=93069430; PubMed=9852099;  
Liu X., Hwang H., Cao L., Wen D., Liu N., Graham R.M., Zhou M.;  
"Release of the neuregulin functional polypeptide requires its cytoplasmic tail.";  
J Biol. Chem. 273:34335-34340(1998).  
[5]  
INTERACTION WITH LIMK1.  
MEDLINE=98352096; PubMed=9685409;  
Wang J.Y., Frenzel K.E., Wen D., Falls D.L.;  
"Transmembrane neuregulins interact with Lim kinase 1, a cytoplasmic protein kinase implicated in development of visuospatial cognition.";  
J. Biol. Chem. 273:20523-20534(1998).  
-1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORCEPTORS, RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND ACTIVATION OF THE ERBB RECEPTORS. THE MULTIPLE ISOFORMS PERFORM DIVERSE FUNCTIONS SUCH AS INDUCING GROWTH AND DIFFERENTIATION OF EPITHELIAL, GLIAL, NEURONAL, AND SKELETAL MUSCLE CELLS; INDUCING EXPRESSION OF ACETYLCHOLINE RECEPTOR IN SYNAPTIC VESICLES DURING THE FORMATION OF THE NEUROMUSCULAR JUNCTION; STIMULATING LOBULOBLAR BUDDING AND MILK PRODUCTION IN THE MAMMARY GLAND AND INDUCING DIFFERENTIATION OF MAMMARY TUMOR CELLS; STIMULATING

CC SCHEMANN CELL PROLIFERATION; IMPLICATION IN THE DEVELOPMENT OF THE  
CC MYOCARDIUM SUCH AS TRABECULATION OF THE DEVELOPING HEART (BY  
CC SIMILARITY).  
CC -1- SUBUNIT: THE CYTOPLASMIC DOMAIN INTERACTS WITH THE LIM DOMAIN  
CC REGION OF LIMK1.  
CC -1- SUBCELLULAR LOCATION: EXISTS AS A TYPE I MEMBRANE PROTEIN AND AS A  
CC PROTOTYPICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-  
CC BOUND FORM DOES NOT SEEM TO BE ACTIVE.  
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 8 ISOFORMS: ALPHA2A/NDP38,  
CC ALPHA2B/NDP19, ALPHA2C/NDP44, BETA1, BETA2/NDP40, BETA2A/NDP22,  
CC BETA3/NDP4 AND BETA4/NDP42A (SHOWN HERE); ARE PRODUCED BY  
CC ALTERNATIVE SPLICING.  
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED. MOST TISSUES CONTAIN ALPHA2A  
CC AND ALPHA2B ISOFORMS. ALPHA2A AND BETA2 ARE THE PREDOMINANT FORMS  
CC IN MESENCHYMAL AND NONNEURONAL ORGANS. BETA1 IS ENRICHED IN BRAIN  
CC AND SPINAL CORD, BUT NOT IN MUSCLE AND HEART. ALPHA2C IS HIGHLY  
CC EXPRESSED IN SPINAL CORD, MODERATELY IN LUNG, BRAIN OVARY, AND  
CC STOMACH, IN LOW AMOUNTS IN THE KIDNEY, SKIN AND HEART AND NOT  
CC DETECTED IN THE LIVER, SPLEEN, AND PLACENTA.  
CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION OF  
CC TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE  
CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN  
CC DIMERIZATION.  
CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE  
CC DOMAIN.  
CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE  
CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR  
CC FORM.  
CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE.  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.  
CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.  
CC  
CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION  
CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION  
CC AT THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS  
CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY  
CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL  
CC ENTITIES REQUIRES A LICENSE AGREEMENT (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: U02315; AAA19940.1; -  
CC EMBL: U02316; AAA19941.1; -  
CC EMBL: U02317; AAA19942.1; -  
CC EMBL: U02318; AAA19943.1; -  
CC EMBL: U02319; AAA19944.1; -  
CC EMBL: U02320; AAA19945.1; -  
CC EMBL: U02321; AAA19946.1; -  
CC EMBL: U02322; AAA19947.1; -  
CC EMBL: U02323; AAA19948.1; -  
CC EMBL: U02324; AAA19949.1; -  
CC EMBL: M92430; -; NOT\_ANNOTATED\_CDS.  
CC HSP: Q12784; 1HRP.  
CC InterPro: IPR000561; EGF-like.  
CC InterPro: IPR003006; Ig\_MHC.  
CC InterPro: IPR003598; Ig\_C2.  
CC InterPro: IPR002154; Neuregulin.  
CC Pfam: PF00008; EGF.1.  
CC Pfam: PF00047; Ig.1.  
CC Pfam: PF02158; Neuregulin.1.  
CC PRINTS: PR01089; NEUREGULIN.  
CC SMART: SM00181; EGF.1.  
CC SMART: SM00408; IGC2.1.  
CC PROSITE: PS00022; EGF\_1.1.  
CC PROSITE: PS01186; EGF\_2.1.  
CC Growth factor: EGF-like domain; Immunoglobulin domain; glycoprotein;  
CC Transmembrane; Multigene family; Alternative splicing.  
CC PROPEP 13  
CC CHAIN 14 662 PRO-NEUREGULIN-1, MEMBRANE-BOUND FORM.  
CC CHAIN 14 264 NEUREGULIN-1.  
CC DOMAIN 14 265 EXTRACELLULAR (POTENTIAL).  
CC TRANSMEM 266 288 INTERNAL SIGNAL SEQUENCE (POTENTIAL).  
CC DOMAIN 289 662 CYTOPLASMIC (POTENTIAL).





FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 1 127 MATSSEPLDYSIAPOTDWNSSIVPPPLKEMKQEVAV  
 FT VARSPLIC 1 127 GQVLVLCETTSYSPALRFKWLKNGKITKKNPENKIK  
 FT VARSPLIC 1 127 KOKYSELHIYRATLADAGEACRVSKLNDSTKASVIT  
 FT VARSPLIC 1 127 DTNA -> MSEVGETEPSPSAOLSPASISGGPAERNMG  
 FT VARSPLIC 1 127 PHREDSRVGVAGLASCVCYCLEAERKIGKLNSEKICIAI  
 FT VARSPLIC 1 127 LACLSICICINACIKVYFVDKITEYSDSPHLDGIGODR  
 FT VARSPLIC 1 127 STVDPTLALSAWSEVYASFPPIPSLESKREVYQDSLY  
 FT VARSPLIC 1 127 PSRRPLQPSLYNRLIDVGLMSATPSLSPSELEPTASQ  
 FT VARSPLIC 1 127 ATEINLOTAPKLS (IN ISOFORM BETA1A, ISOFORM  
 FT VARSPLIC 1 127 BETA2A AND ISOFORM BETA2B).  
 FT VARSPLIC 1 127 MISSING (IN ISOFORM BETA2A AND ISOFORM  
 FT VARSPLIC 1 127 BETA2B).  
 FT VARSPLIC 1 127 VSATPTPARNSPVDFHTP -> HTPPTSLLAGKVSILRV  
 FT VARSPLIC 1 127 (IN ISOFORM BETA2B).  
 FT VARSPLIC 1 127 MISSING (IN ISOFORM BETA2B).  
 SQ SEQUENCE 602 AA; 67453 MW; 41830E56CE5D346 CRC64;  
 Query Match 35.68; Score 98.5; DB 1; Length 602;  
 Best Local Similarity 33.38; Pred. No. 0.00014;  
 Matches 16; Conservative 11; Mismatches 20; Indels 1; Gaps 1;  
 Oy 1 HFKPCDKDLAYCNDCEPVITLGSNKH-CRCKEGYGVRCDFL 47  
 Db 137 HLTCKDIKOKAFVNGSECTVMDLPNPPYLCRCRPNFEGRCQNTY 184  
 RESULT 10  
 NR2\_HUMAN STANDARD; PRT; 850 AA.  
 AC 014511;  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Pro-neuregulin-2 precursor (Pro-NRG2) [containing: Neuregulin-2 (NRG-2)  
 DE (Neural-and thymus-derived activator for ERB B kinases) (NTAK)  
 DE (Divergent of neuregulin 1) (DON-1)].  
 DE NRG2 OR NTAK.  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 CC NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RP TISSUE-Neuroblastoma;  
 RX MEDLINE-98006324; PubMed-9348101;  
 RA Higashiyama S., Horikawa M., Yamada K., Ichino N., Nakano N.,  
 RA Nakagawa T., Miyagawa J., Matsushita N., Nagatsu T., Taniguchi N.,  
 RA Ishiguro H.;  
 RA "A novel brain-derived member of the epidermal growth factor family  
 RA that interacts with ErbB3 and ErbB4.";  
 RT J. Biochem. 122:675-680(1997).  
 RL [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS DON-1B AND DON-1R).  
 RP TISSUE-Fetal brain;  
 RX MEDLINE-97342638; PubMed-9199335;  
 RA Busfield S.J., Michnick D.A., Chickering T.W., Revett T.L., Ma J.,  
 RA Woolf E.A., Comrack C.A., Dussault B.J., Woolf J., Gooderli A.D.J.,  
 RA Gearing D.P.;  
 RA "Characterization of a neuregulin-related gene, Don-1, that is highly  
 RA expressed in restricted regions of the cerebellum and hippocampus.";  
 RT Mol. Cell. Biol. 17:4007-4014(1997).  
 RL [3]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5 AND 6).  
 RP TISSUE-Lung, and Fetal brain;  
 RX MEDLINE-99295836; PubMed-10369162;  
 RA Ring H.Z., Chang H., Gullbot A., Brice A., Leguenn E., Franke U.;  
 RA "The human neuregulin 2 (NRG2) gene: cloning, mapping and evaluation  
 RA as a candidate for the autosomal recessive form of Charcot-Marie-Tooth  
 RT disease linked to 5q.";

RL Hum. Genet. 104:326-332(1999).  
 CC -1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE  
 CC RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,  
 CC RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND  
 CC ACTIVATION OF THE ERBB RECEPTORS. MAY ALSO PROMOTE THE  
 CC HETERO-DIMERIZATION WITH THE EGF RECEPTOR.  
 CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS  
 CC A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE  
 CC MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).  
 CC -1- ALTERNATIVE PRODUCTS: 8 ISOFORMS; 1 (SHOWN HERE), 2, 3, 4, 5, 6,  
 CC DON-1B AND DON-1R. ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: RESTRICTED TO THE CEREBELLUM IN THE ADULT.  
 CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION  
 CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE  
 CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN  
 CC DIMERIZATION (BY SIMILARITY).  
 CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE  
 CC DOMAIN (BY SIMILARITY).  
 CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE BY THE EGF-LIKE  
 CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR  
 CC FORM (BY SIMILARITY).  
 CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sb.ch](mailto:license@sib-sb.ch)).  
 EMBL AB005060. BAA23417.1. -  
 EMBL AF119162. AAF28848.1. -  
 EMBL AF119151. AAF28848.1. JOINED.  
 EMBL AF119152. AAF28848.1. JOINED.  
 EMBL AF119153. AAF28848.1. JOINED.  
 EMBL AF119154. AAF28848.1. JOINED.  
 EMBL AF119155. AAF28848.1. JOINED.  
 EMBL AF119156. AAF28848.1. JOINED.  
 EMBL AF119157. AAF28848.1. JOINED.  
 EMBL AF119158. AAF28848.1. JOINED.  
 EMBL AF119159. AAF28848.1. JOINED.  
 EMBL AF119160. AAF28848.1. JOINED.  
 EMBL AF119161. AAF28848.1. JOINED.  
 EMBL AF119162. AAF28848.1. JOINED.  
 EMBL AF119163. AAF28848.1. JOINED.  
 EMBL AF119164. AAF28848.1. JOINED.  
 EMBL AF119165. AAF28848.1. JOINED.  
 EMBL AF119166. AAF28848.1. JOINED.  
 EMBL AF119167. AAF28848.1. JOINED.  
 EMBL AF119168. AAF28848.1. JOINED.  
 EMBL AF119169. AAF28848.1. JOINED.  
 EMBL AF119170. AAF28848.1. JOINED.  
 EMBL AF119171. AAF28848.1. JOINED.  
 EMBL AF119172. AAF28848.1. JOINED.  
 EMBL AF119173. AAF28848.1. JOINED.  
 EMBL AF119174. AAF28848.1. JOINED.  
 EMBL AF119175. AAF28848.1. JOINED.  
 EMBL AF119176. AAF28848.1. JOINED.  
 EMBL AF119177. AAF28848.1. JOINED.  
 EMBL AF119178. AAF28848.1. JOINED.  
 EMBL AF119179. AAF28848.1. JOINED.  
 EMBL AF119180. AAF28848.1. JOINED.  
 EMBL AF119181. AAF28848.1. JOINED.  
 EMBL AF119182. AAF28848.1. JOINED.  
 EMBL AF119183. AAF28848.1. JOINED.  
 EMBL AF119184. AAF28848.1. JOINED.  
 EMBL AF119185. AAF28848.1. JOINED.  
 EMBL AF119186. AAF28848.1. JOINED.  
 EMBL AF119187. AAF28848.1. JOINED.  
 EMBL AF119188. AAF28848.1. JOINED.  
 EMBL AF119189. AAF28848.1. JOINED.  
 EMBL AF119190. AAF28848.1. JOINED.  
 EMBL AF119191. AAF28848.1. JOINED.  
 EMBL AF119192. AAF28848.1. JOINED.  
 EMBL AF119193. AAF28848.1. JOINED.  
 EMBL AF119194. AAF28848.1. JOINED.  
 EMBL AF119195. AAF28848.1. JOINED.  
 EMBL AF119196. AAF28848.1. JOINED.  
 EMBL AF119197. AAF28848.1. JOINED.  
 EMBL AF119198. AAF28848.1. JOINED.  
 EMBL AF119199. AAF28848.1. JOINED.  
 EMBL AF120000. AAF28848.1. JOINED.



```

CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
CC DIMERIZATION (BY SIMILARITY).
CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
CC DOMAIN (BY SIMILARITY).
CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
CC FORM (BY SIMILARITY).
CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
CC SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D89995; BAA23344.1; -
CC DR EMBL: D89996; BAA23345.1; -
CC DR EMBL: D89997; BAA23346.1; -
CC DR EMBL: D89998; BAA23347.1; -
CC DR EMBL: AB001576; BAA23348.1; -
CC DR HSSP: Q12784; 1HRE.
CC DR InterPro: IPR000561; EGF-like.
CC DR InterPro: IPR003006; IG-MHC.
CC DR InterPro: IPR003598; IG-C2.
CC DR InterPro: IPR002154; Neuregulin.
CC DR Pfam: PF00047; 1g; 1.
CC DR Pfam: PF02158; Neuregulin; 2.
CC DR SMART: SM00181; EGF; 1.
CC DR SMART: SM00408; IGC2; 1.
CC DR PROSITE: PS00022; EGF_1; 1.
CC DR PROSITE: PS01186; EGF_2; 1.
CC DR Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein;
CC Transmembrane; Multigene family; Alternative splicing.
CC PROPEP 127
CC CHAIN 128 868 PRO-NEUREGULIN-2, MEMBRANE-BOUND FORM.
CC DOMAIN 128 428 NEUREGULIN-2.
CC TRANSSEM 430 450 EXTRACELLULAR (POTENTIAL).
CC DOMAIN 451 868 INTERNAL SIGNAL SEQUENCE (POTENTIAL).
CC DOMAIN 266 334 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 346 356 IG-LIKE C2-TYPE DOMAIN.
CC DOMAIN 357 398 SER/THR-RICH.
CC DOMAIN 357 398 EGF-LIKE.
CC DOMAIN 22 32 POLY-SER.
CC DOMAIN 35 45 POLY-SER.
CC DOMAIN 56 59 POLY-SER.
CC DOMAIN 103 106 POLY-THR.
CC DOMAIN 739 745 POLY-ALA.
CC DISULFID 273 745 POLY-PRO.
CC DISULFID 327 745 BY SIMILARITY.
CC DISULFID 361 375 BY SIMILARITY.
CC DISULFID 369 386 BY SIMILARITY.
CC DISULFID 388 397 BY SIMILARITY.
CC CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 362 362 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC VASPLIC 1 108 MISSING (IN ISOFORM NR2-BETA).
CC VASPLIC 220 222 PLY -> FFF (IN ISOFORM NTAK-ALPHA2-1P).
CC VASPLIC 222 222 C-> G (IN ISOFORM NTAK-GAMMA).
CC VASPLIC 388 388 MISSING (IN ISOFORM NTAK-GAMMA).
CC VASPLIC 389 868 NGFGORCLEPLRLVMDPKDKHGLSELEKE -> VGYTGDCQGFAMV
CC VASPLIC 390 412 NFS (IN ISOFORM NR2-BETA).
CC VASPLIC 390 868 NGFGORCLEPLRLVMDPKDKHGLSELEKE -> VGYTG
CC VASPLIC 390 868 DRCQGFAMVNSK (IN ISOFORM NTAK-BETA).
CC VASPLIC 390 421

```

```

FT VASPLIC 414 421 MISSING (IN ISOFORM NTAK-ALPHA2 AND
FT VASPLIC 414 439 ISOFORM NTAK-ALPHA2B).
FT VASPLIC 440 868 HGFELKEAEELYQKRLITGICVA -> SYLMDPTGTV
FT VASPLIC 440 868 SSSQSTSPSTLDN (IN ISOFORM NR2-ALPHA).
FT CONFLICT 117 117 MISSING (IN ISOFORM NR2-ALPHA).
FT CONFLICT 117 117 S -> F (IN REF. 2).
FT CONFLICT 724 724 R -> H (IN REF. 2).
SQ SEQUENCE 868 AA; 93776 MW; 3C7D4D94DBE64DE2 CRC64;

Query Match 32.5%; Score 90; DB 1; Length 868;
Best Local Similarity 34.9%; Pred. No. 0.002;
Matches 15; Conservative 9; Mismatches 17; Indels 2; Gaps 1;

QY 1 HFPCRDKDLAYCLNDGCEVFETLNGSHKCRKRGYGVRC 43
DB 357 HAKCKETAKSYCVNGVCIYIGT--NQLSCPKMFGQRQRC 357

RESULT 12
ID EREG_HUMAN STANDARD; PRT; 169 AA.
AC 014944;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Eregulin precursor.
GN EREG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC TISSUE=Colorectal adenocarcinoma;
RX MEDLINE=97479200; PubMed=9337852;
RA Toyoda H., Komurasaki T., Uchida D., Morimoto S.;
RT "Distribution of mRNA for human eregulin, a differentially expressed
RT member of the epidermal growth factor family.";
RL Biochem. J. 326:69-75(1997).
CC -1- FUNCTION: MAY BE A MEDIATOR OF LOCALIZED CELL PROLIFERATION. AS A
CC MITOGEN IT MAY STIMULATE CELL PROLIFERATION AND/OR ANGIOGENESIS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (PRECURSOR FORM).
CC -1- EXTRACELLULAR (MATURE FORM).
CC -1- TISSUE SPECIFICITY: IN NORMAL ADULTS EXPRESSED PREDOMINANTLY IN THE
CC PLACENTA AND PERIPHERAL BLOOD LEUCOCYTES. HIGH LEVELS WERE DETECTED
CC IN CARCINOMAS OF THE BLADDER, LUNG, KIDNEY AND COLON.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D30783; BAA22146.1; -
CC MIM: 602061; -
CC InterPro: IPR000561; EGF-like.
CC InterPro: IPR001336; EGF_1.
CC PRINTS: PR00009; EGF_TGF.
CC SMART: SM00181; EGF; 1.
CC PROSITE: PS01186; EGF_2; 1.
CC Growth factor; Mitogen; Glycoprotein; EGF-like domain; Transmembrane;
CC Signal.
FT SIGNAL 1 29
FT PROPEP 30 59 POTENTIAL.
FT CHAIN 60 108 POTENTIAL.
FT PROPEP 109 169 EPIREGULIN.
FT DOMAIN 60 119 REMOVED IN MATURE FORM (BY SIMILARITY).
FT TRANSSEM 120 140 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 141 169 POTENTIAL.
FT DOMAIN 141 169 CYTOPLASMIC (POTENTIAL).

```

FT DOMAIN 145 152 ARG/LYS-RICH (BASIC).  
 FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT DOMAIN 64 104 EGF-LIKE.  
 FT DISULFID 68 81 BY SIMILARITY.  
 FT DISULFID 76 92 BY SIMILARITY.  
 FT DISULFID 94 103 BY SIMILARITY.  
 SQ SEQUENCE 169 AA; 19044 MW; 17F3926ADFB2BDEE CAC64;

Query Match 32.3%; Score 89.5; DB 1; Length 169;  
 Best Local Similarity 40.5%; Pred. No. 0.0005;  
 Matches 17; Conservative 7; Mismatches 15; Indels 3; Gaps 2;

Oy 5 CRDRLAYCNDGECFVETLTGSHKRCCKEGTQVRCDOF 46  
 Db 68 CSSDMNGYCLH-GOCITYLVM--SONYCRCEVGYGVACEHF 106

## RESULT 13

GRA\_SFVKA STANDARD; PRT; 80 AA.  
 ID GRA\_SFVKA STANDARD; PRT; 80 AA.  
 AC P08441;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Growth factor.  
 OS Shope fibroma virus (strain Kasza) (SFV).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Leporipoxvirus.  
 OX NCBI\_TaxID=10272;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8712751; PubMed=3031480;  
 RT Chang W., Opton C., Hu S.-L., Purchio A.F., McFadden G.;  
 "The genome of Shope fibroma virus, a tumorigenic poxvirus, contains  
 a growth factor gene with sequence similarity to those encoding  
 epidermal growth factor and transforming growth factor alpha.";  
 RL Mol. Cell Biol. 7:535-540(1987).  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb-sdb.ch](mailto:license@sdb-sdb.ch)).  
 CC EMBL, M15921; AAA66873.1; -  
 DR PIR: A26723; EGVZSF.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR001336; EGF\_1.  
 DR PRINTS: PR00009; EGF\_TGF.  
 DR SMART: SM00181; EGF\_1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; FALSE\_NEG.  
 KV EGF-like domain; growth factor; glycoprotein.  
 FT DOMAIN 29 73  
 FT DISULFID 33 47 EGF-LIKE.  
 FT DISULFID 41 61 BY SIMILARITY.  
 FT DISULFID 63 72 BY SIMILARITY.  
 FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 80 AA; 9210 MW; C48D30878D2ED58 CRC64;

Query Match 30.7%; Score 85; DB 1; Length 80;  
 Best Local Similarity 39.2%; Pred. No. 0.00087;  
 Matches 20; Conservative 6; Mismatches 17; Indels 8; Gaps 3;

Oy 1 HFKCRDKLAYCNDGECFVETLTGSHKRCCKEGTQVRCDOF 47  
 Db 29 HAKVCNHDYENYCLNNGCFIALDNVST--PFCVCRIYNEGSRG-QFI 75

RESULT 14  
 NCCL\_MOUSE STANDARD; PRT; 2531 AA.  
 ID NCCL\_MOUSE  
 AC Q01705;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Neurogenic locus notch homolog protein 1 precursor (NOTCH protein).  
 GN NOTCH1 OR NOTCH.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=93194170; PubMed=8449489.  
 RA Franco del Amo F., Gendron-Maguire M., Swiatek P.J., Jenkins N.A.,  
 RA Copeland N.G., Gridley T.;  
 RT "Cloning, analysis, and chromosomal localization of Notch-1, a mouse  
 RT homolog of Drosophila Notch.";  
 RL Genomics 15:259-264(1993).  
 RN [2]  
 RP SEQUENCE OF 1551-2170 FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=93048835; PubMed=1425352;  
 RA Franco del Amo F., Smith D.E., Swiatek P.J., Gendron-Maguire M.,  
 RA Greenspan R.J., McMahon A.P., Gridley T.;  
 RT "Expression pattern of Notch, a mouse homolog of Drosophila Notch,  
 RT suggests an important role in early postimplantation mouse  
 RT development.";  
 RL Development 115:737-744(1992).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED ALMOST UNIFORMLY IN EARLY EMBRYOS.  
 CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.  
 CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.  
 CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb-sdb.ch](mailto:license@sdb-sdb.ch)).  
 CC EMBL, Z11886; CAA77941.1; -  
 DR HSSP: P00740; IEDM.  
 DR MGD: MGI:97363; Notch1.  
 DR InterPro: IPR002110; ANK.  
 DR InterPro: IPR000152; ASX\_hydroxyl.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR00742; EGF\_2.  
 DR InterPro: IPR001881; EGF\_CA.  
 DR InterPro: IPR001438; EGF\_II.  
 DR InterPro: IPR000800; Notch.  
 DR Pfam: PF00023; ank; 6.  
 DR Pfam: PF00008; EGF; 35.  
 DR Pfam: PF00066; notch; 3.  
 DR PRINTS: PR00010; EGFBL00D.  
 DR PRINTS: PR01452; NOTCH.  
 DR SMART: SM00248; ANK; 3.  
 DR SMART: SM00179; EGF\_CA; 23.  
 DR SMART: SM00001; EGF\_LIKE; 11.  
 DR SMART: SM00004; NL; 2.  
 DR PROSITE: PS50088; ANK\_REPEAT; 2.  
 DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 22.  
 DR PROSITE: PS00022; EGF\_1; 34.  
 DR PROSITE: PS01186; EGF\_2; 27.

FT	DISULFID	299	312	BY SIMILARITY.
FT	DISULFID	306	321	BY SIMILARITY.
FT	DISULFID	323	332	BY SIMILARITY.
FT	DISULFID	339	350	BY SIMILARITY.
FT	DISULFID	344	359	BY SIMILARITY.
FT	DISULFID	361	370	BY SIMILARITY.
FT	DISULFID	376	387	BY SIMILARITY.
FT	DISULFID	381	398	BY SIMILARITY.
FT	DISULFID	400	409	BY SIMILARITY.
FT	DISULFID	416	429	BY SIMILARITY.
FT	DISULFID	423	438	BY SIMILARITY.
FT	DISULFID	440	449	BY SIMILARITY.
FT	DISULFID	456	467	BY SIMILARITY.
FT	DISULFID	478	487	BY SIMILARITY.
FT	DISULFID	494	505	BY SIMILARITY.
FT	DISULFID	499	514	BY SIMILARITY.
FT	DISULFID	516	525	BY SIMILARITY.
FT	DISULFID	532	543	BY SIMILARITY.
FT	DISULFID	537	552	BY SIMILARITY.
FT	DISULFID	554	563	BY SIMILARITY.
FT	DISULFID	570	580	BY SIMILARITY.
FT	DISULFID	575	589	BY SIMILARITY.
FT	DISULFID	591	600	BY SIMILARITY.
FT	DISULFID	607	618	BY SIMILARITY.
FT	DISULFID	612	627	BY SIMILARITY.
FT	DISULFID	629	638	BY SIMILARITY.
FT	DISULFID	645	655	BY SIMILARITY.
FT	DISULFID	650	664	BY SIMILARITY.
FT	DISULFID	666	675	BY SIMILARITY.
FT	DISULFID	682	693	BY SIMILARITY.
FT	DISULFID	687	702	BY SIMILARITY.
FT	DISULFID	704	713	BY SIMILARITY.
FT	DISULFID	720	730	BY SIMILARITY.
FT	DISULFID	725	739	BY SIMILARITY.
FT	DISULFID	741	750	BY SIMILARITY.
FT	DISULFID	757	768	BY SIMILARITY.
FT	DISULFID	762	777	BY SIMILARITY.
FT	DISULFID	779	788	BY SIMILARITY.
FT	DISULFID	795	806	BY SIMILARITY.
FT	DISULFID	800	815	BY SIMILARITY.
FT	DISULFID	817	826	BY SIMILARITY.
FT	DISULFID	833	844	BY SIMILARITY.
FT	DISULFID	857	866	BY SIMILARITY.
FT	DISULFID	873	884	BY SIMILARITY.
FT	DISULFID	878	883	BY SIMILARITY.
FT	DISULFID	895	904	BY SIMILARITY.
FT	DISULFID	911	922	BY SIMILARITY.
FT	DISULFID	916	931	BY SIMILARITY.
FT	DISULFID	933	942	BY SIMILARITY.
FT	DISULFID	987	998	BY SIMILARITY.
FT	DISULFID	992	1007	BY SIMILARITY.
FT	DISULFID	1009	1018	BY SIMILARITY.
FT	DISULFID	1025	1036	BY SIMILARITY.
FT	DISULFID	1030	1045	BY SIMILARITY.
FT	DISULFID	1047	1056	BY SIMILARITY.
FT	DISULFID	1063	1074	BY SIMILARITY.

Query Match 30.0%; Score 83; DB 1; Length 2531;  
Best Local Similarity 41.7%; Pred. No. 0.037;  
Matches 15; Conservative 4; Mismatches 13; Indels 4; Gaps 1

9 DLAFLNDGCEVITLGSRRHCKCKEYQGVCD 44

Db 1064 DSAFCNGKRCW---QNTQYHCECRSGMTGVNCD 1095

RESULT 15  
BTRC\_MOUSE  
AC 005928;  
01-FEB-1994 (Rel. 28, Created)  
STANDARD; PRT; 177 AA.

DT 01-FEB-1994 (Rel. 28, last sequence update)  
DT 16-OCT-2001 (Rel. 40, last annotation update)  
DE Betacellulin precursor (BTC).  
GN BTC OR BCN.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-54; 64-71 AND 75-111.  
RC TISSUE=pancreas;  
RX MEDLINE=93206093; PubMed=8456283;  
RA Shing Y., Christofori G., Hanahan D., Ono Y., Sasada R.,  
RA Igarashi K., Folkman J.;  
RT "Betacellulin: a mitogen from pancreatic beta cell tumors.";  
RL Science 259:1604-1607 (1993).  
CC -1- FUNCTION: POTENT MITOGEN FOR RETINAL PIGMENT EPITHELIAL CELLS  
CC AND VASCULAR SMOOTH MUSCLE CELLS. THE EFFECTS OF BETACELLULIN  
CC ARE PROBABLY MEDIATED BY THE EGF RECEPTOR AND OTHER RELATED  
CC RECEPTORS.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (PRECURSOR FORM);  
CC EXTRACELLULAR (MATURE FORM).  
CC -1- TISSUE SPECIFICITY: FOUND IN SEVERAL MOUSE TISSUES INCLUDING  
CC KIDNEY, UTERUS AND LIVER AS WELL AS IN BETA TUMOR CELL LINE AND  
CC MCF-7 CELLS. IT IS NOT DETECTED IN THE BRAIN.  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: L08394; AAA40511.1; .  
DR PIR: A37408; A37408.  
DR HSSP: P01135; 3TGF.  
DR MGD: MGI:99439; Btc.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR001336; EGF\_1.  
DR Pfam: PF00008; EGF; 1.  
DR PRINTS: PR00009; EGFTEGF.  
DR SMART: SM00181; EGF; 1.  
DR SMART: SM00001; EGF-like; 1.  
DR PROSITE: PS00022; EGF\_1; 1.  
DR PROSITE: PS01186; EGF\_2; 1.  
DR Growth factor; Mitogen; Glycoprotein; EGF-like domain; Transmembrane;  
KW signal.  
FT 1 31  
FT CHAIN 32 111  
FT PROPEP 112 177  
FT DOMAIN 32 118  
FT TRANSLEM 119 139  
FT DOMAIN 140 177  
FT DOMAIN 65 105  
FT DOMAIN 146 153  
FT DISULFID 69 82  
FT DISULFID 77 93  
FT DISULFID 95 104  
FT CARBOHYD 34 42  
FT CARBOHYD 42 42  
FT CARBOHYD 52 52  
SO SEQUENCE 177 AA; 19664 MW; 066B34F0E13F82B CRC64;

Query Match 29.8%; Score 82.5; DB 1; Length 177;  
Best Local Similarity 39.1%; Pred. No. 0.0036;  
Matches 18; Conservative 9; Mismatches 14; Indels 5; Gaps 3;

QY 1 HFKPCKDLAYCLINDSEC-FVETITLGSKHKCKEGYGVRCQ 45  
DB 65 HFSRCPRQYKHYCIH-GRCRFVDEQTPS---CICEKGYGARCER 106

Search completed: September 13, 2002, 08:55:43  
Job time: 350 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 08:44:43 ; Search time 18.37 Seconds  
(without alignments)  
245.846 Million cell updates/sec

Title: US-09-877-665-4

Perfect score: 277  
Sequence: 1 HFKPCRDKDLAYCLNDECFYETLTGSHKRCCKEGYGVRCDOFL 47

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: PIR 71:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	277	100.0	713	2	T44447
2	113.5	41.0	125	2	I38405
3	113.5	41.0	462	2	I38405
4	113.5	41.0	640	2	A43273
5	111.5	40.3	639	2	I61719
6	105.5	38.1	125	2	S62676
7	104.5	37.7	175	2	I38408
8	104.5	37.7	241	2	S32359
9	104.5	37.7	241	2	D43273
10	104.5	37.7	296	2	A56943
11	104.5	37.7	422	2	S32357
12	104.5	37.7	637	2	C43273
13	104.5	37.7	645	2	B43273
14	102.5	37.0	230	2	A56210
15	102.5	37.0	636	2	I61718
16	102.5	37.0	662	2	I61722
17	98.5	35.6	602	2	A43769
18	92.5	33.4	2180	2	T29764
19	90	32.5	850	2	JC5700
20	90	32.5	860	2	JC5702
21	90	32.5	868	2	JC5701
22	85	30.7	80	1	EGVZSF
23	84.5	30.5	1220	2	A56136
24	83	30.0	46	2	JT0747
25	83	30.0	162	2	S68401
26	83	30.0	861	2	A48825
27	83	30.0	2531	2	A46019
28	82.5	29.8	177	2	A37408
29	82	29.6	85	1	EGVZM1

30	82	29.6	230	2	A44074	probable EGF-like
31	82	29.6	264	2	T22380	hypothetical prote
32	80.5	29.1	178	2	JC1467	betacellulin precu
33	80	28.9	907	2	T27317	hypothetical prote
34	79.5	28.7	140	1	WMV29	growth factor - va
35	79.5	28.7	140	2	T30766	growth factor - va
36	79.5	28.7	142	1	WMV23C	transforming growt
37	79.5	28.7	159	1	WFRN1	transforming growt
38	79.5	28.7	159	1	S27195	transforming growt
39	79.5	28.7	159	2	I57497	transforming growt
40	79	28.5	2531	2	S18188	transforming growt
41	78	28.2	1207	1	EGH0	transforming growt
42	77.5	28.0	722	2	I48324	transforming growt
43	77.5	28.0	2352	2	T30201	transforming growt
44	76.5	27.6	1372	2	T25933	transforming growt
45	76	27.4	482	2	JC5092	transforming growt

## ALIGNMENTS

RESULT 1  
T44447  
neuregulin-3 [imported] - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000  
C:Accession: T44447  
R:Zhang, D.; Sliwkowski, M.X.; Mart, M.; Frantz, G.; Akita, R.; Sun, Y.; Hillan, K.;  
Proc. Natl. Acad. Sci. U.S.A. 94, 9562-9567, 1997  
A:Title: Neuregulin-3 (NRG3): A novel neural tissue-enriched protein that binds and a  
A:Reference number: Z22773; MUID:97420720  
A:Accession: T44447  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-713 <ZHA>  
A:Cross-references: EMBL:AF010130; NID:92429163; PIDN:AA870914.1; PID:92429164  
C:Genetics:  
A:Gene: NRG3

Query Match 100.0%; Score 277; DB 2; Length 713;  
Best Local Similarity 100.0%; Pred. No. 4.7e-24;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDKDLAYCLNDECFYETLTGSHKRCCKEGYGVRCDOFL 47  
DB 288 HFKPCRDKDLAYCLNDECFYETLTGSHKRCCKEGYGVRCDOFL 334

RESULT 2  
I38405  
neu differentiation factor - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 05-Nov-1999  
C:Accession: I38405  
R:Wen, D.; Suggs, S.V.; Karunakaran, D.; Lau, N.; Cupples, R.L.; Luo, Y.; Janssen, A.  
Mol. Cell. Biol. 14, 1909-1919, 1994  
A:Title: Structural and functional aspects of the multiplicity of Neu differentiation  
A:Reference number: A56210; MUID:94158863  
A:Accession: I38405  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-125 <RES>  
A:Cross-references: EMBL:U02327; NID:9408404; PIDN:AAA19952.1; PID:9408405

Query Match 41.0%; Score 113.5; DB 2; Length 125;  
Best Local Similarity 34.8%; Pred. No. 5.1e-06;  
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;  
OY 1 HFKPCRDKDLAYCLNDECFYETLTGSHKRCCKEGYGVRCDO 45  
DB 56 HFKPCRDKDLAYCLNDECFYETLTGSHKRCCKEGYGVRCDO 101



RESULT 3  
138404  
neu differentiation factor - human  
C:Species: Homo sapiens (man)  
C>Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 05-Nov-1999  
C:Accession: 138404  
R:Wen, D.; Suggs, S.V.; Karunagaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.; Mol. Cell. Biol. 14, 1909-1919, 1994  
A>Title: Structural and functional aspects of the multiplicity of Neu differentiation factor  
A:Reference number: A56210; MUID:94158863  
A:Accession: 138404  
A>Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-462 <RES>  
A:Cross-references: EMBL:U02326; NID:9408402; PIDN:AAA19951.1; PID:9408403

Query Match 41.0%; Score 113.5; DB 2; Length 462;  
Best Local Similarity 34.8%; Pred. No. 1.6e-05;  
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

OY 1 HEKPRDKDLAVCLNDGECFVETLTGSHKH-CRCKEGYGVRCDO 45  
DB 178 HLVKCAEKETFCVNGGECFVWDLNPSRYLCKCPGFGARCTE 223

RESULT 4  
A43273  
heregulin precursor, splice form alpha - human  
N:Alternate names: breast cancer cell differentiation factor p45; Neu differentiation factor  
C:Species: Homo sapiens (man)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 11-Jan-2000  
C:Accession: A43273; A48498; A38155  
R:Holmes, W.E.; Sliwowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yansuri  
Science 259, 1205-1210, 1992  
A>Title: Identification of heregulin, a specific activator of p185(erbB2).  
A:Reference number: A43273; MUID:92271253  
A:Accession: A43273  
A>Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-640 <HOL>  
A:Experimental source: breast tumor cell line, MDA-MB-231, ATCC HTB 26  
A>Note: Sequence extracted from NCBI backbone (NCBI:103250)  
R:Clousson, J.M.; Plozman, G.D.; Carlton, G.W.; Green, J.M.; Shoyab, M.  
J. Biol. Chem. 268, 18407-18410, 1993  
A>Title: Characterization of a breast cancer cell differentiation factor that specifically  
A:Reference number: A48498; MUID:93366731  
A:Accession: A48498  
A:Molecule type: protein  
A:Residues: 20-21, 'X', 23-24, 'XX', 27-28 <CUU>  
R:Peles, E.; Bacus, S.S.; Koshi, R.A.; Lu, H.S.; Wen, D.; Ogden, S.G.; Levy, R.B.; Yarden  
Cell 69, 205-216, 1992  
A>Title: Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein that induce  
A:Reference number: A38155; MUID:92208945  
A:Accession: A38155  
A:Molecule type: protein  
A:Residues: 'X', 15-16, 'X', 18-20, 'RG', 23-24, 'GP', 27, 'E', 29, 'XP', 32-36 <PEL>  
A>Note: sequence extracted from NCBI backbone (NCBI:91347)  
C:Genetics:  
A:Gene: GDB:HGL  
A:Cross-references: GDB:132656; OMIM:142445  
A:Map position: 8p22-8p11  
C:Superfamily: unassigned EGF-related proteins; EGF homology  
C:Keywords: alternative splicing; glycoprotein  
F:182-221/Domain: EGF homology <EGF>

Query Match 41.0%; Score 113.5; DB 2; Length 640;  
Best Local Similarity 34.8%; Pred. No. 2.1e-05;  
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

OY 1 HEKPRDKDLAVCLNDGECFVETLTGSHKH-CRCKEGYGVRCDO 45  
DB 178 HLVKCAEKETFCVNGGECFVWDLNPSRYLCKCPGFGARCTE 223

RESULT 5  
161719  
neu differentiation factor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 01-Dec-2000  
C:Accession: 161719; 161723; 161716; 161717; 161724; A38220  
R:Wen, D.; Suggs, S.V.; Karunagaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.  
Mol. Cell. Biol. 14, 1909-1919, 1994  
A>Title: Structural and functional aspects of the multiplicity of Neu differentiation  
A:Reference number: A56210; MUID:94158863  
A:Accession: 161719  
A>Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-639 <RES>  
A:Cross-references: EMBL:U02319; NID:9408388; PIDN:AAA19944.1; PID:9408389  
A:Accession: 161723  
A>Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-639 <RES>  
A:Cross-references: EMBL:U02323; NID:9408396; PIDN:AAA19948.1; PID:9408397  
A:Accession: 161716  
A>Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-422, 'H', 'NL', 637-638, 'ELRNKAYRSKCMQIOLSATHLRPSITHLGL', <RE4>  
A:Cross-references: EMBL:U02316; NID:9408384; PIDN:AAA19942.1; PID:9408385  
A:Accession: 161717  
A>Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-422, 'H', 'NL', 637-638, 'ELRNKAYRSKCMQIOLSATHLRPSITHLGL', <RE4>  
A:Cross-references: EMBL:U02317; NID:9408384; PIDN:AAA19942.1; PID:9408385  
A:Accession: 161724  
A>Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-422 <RES>  
A:Cross-references: EMBL:U02324; NID:9408398; PIDN:AAA19949.1; PID:9408399  
R:Wen, D.; Peles, E.; Cupples, R.; Suggs, S.V.; Bacus, S.S.; Luo, Y.; Trill, G.; Hu,  
Cell 69, 559-572, 1992  
A>Title: Neu differentiation factor: a transmembrane glycoprotein containing an EGF  
A:Reference number: A38220; MUID:92257596  
A:Accession: A38220  
A>Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-422 <WEN>  
A>Note: sequence extracted from NCBI backbone (NCBI:101767, NCBI:101768)  
C:Superfamily: unassigned EGF-related proteins; EGF homology

Query Match 40.3%; Score 111.5; DB 2; Length 639;  
Best Local Similarity 34.8%; Pred. No. 3.5e-05;  
Matches 16; Conservative 13; Mismatches 16; Indels 1; Gaps 1;

OY 1 HEKPRDKDLAVCLNDGECFVETLTGSHKH-CRCKEGYGVRCDO 45  
DB 178 HLVKCAEKETFCVNGGECFVWDLNPSRYLCKCPGFGARCTE 223

RESULT 6  
S62676  
heregulin isoform alpha 2 - human (fragments)  
N:Alternate names: differentiation factor neu isoform alpha 2  
C:Species: Homo sapiens (man)  
C>Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 13-Mar-1997  
C:Accession: S62676  
R:Hara, S.; Liu, N.; Meng, S.Y.; Lu, H.S.  
Biochim. Biophys. Acta 1292, 168-176, 1996  
A>Title: Isolation and structural characterization of recombinant human neu different  
A:Reference number: S62676; MUID:96139341  
A:Accession: S62676

A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-6;7-16;17-30;31-38;39-58;59-92;93-120;121-125 <MAR>  
C:Keywords: proto-oncogene

Query Match 38.1%; Score 105.5; DB 2; Length 125;  
Best Local Similarity 32.6%; Pred. No. 4.2e-05;  
Matches 15; Conservative 14; Mismatches 16; Indels 1; Gaps 1;  
OY 1 HFKPRDKDLAYCLNDGECFVETLTGSHKH-CRCKEGYGVRCDO 45  
DB 75 HLVKCAEKERTFCVNGGECFVWKDLSNPSRYLCKPCPNFTGDRCONY 120

RESULT 7  
138408  
neu differentiation factor - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 11-Jan-2000  
C:Accession: I38408  
R:Men, D.; Suggs, S.V.; Katunaga, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.; Mol. Cell. Biol. 14, 1909-1919, 1994  
A:Title: Structural and functional aspects of the multiplicity of Neu differentiation factor  
A:Reference number: A56210; MUID:94158863  
A:Accession: I38408  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-175 <RES>  
A:Cross-references: EMBL:U02330; NID:9408410; PIDN:AAA19955.1; PID:9408411  
C:Superfamily: unassigned EGF-related proteins; EGF homology  
F:116-135/Domain: EGF homology <EGF>

Query Match 37.7%; Score 104.5; DB 2; Length 175;  
Best Local Similarity 31.2%; Pred. No. 7.3e-05;  
Matches 15; Conservative 14; Mismatches 18; Indels 1; Gaps 1;  
OY 1 HFKPRDKDLAYCLNDGECFVETLTGSHKH-CRCKEGYGVRCDOFL 47  
DB 112 HLVKCAEKERTFCVNGGECFVWKDLSNPSRYLCKPCPNFTGDRCONY 159

RESULT 8  
532359  
glial growth factor - bovine  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 19-Mar-1997 #sequence\_revision 01-Aug-1997 #text\_change 11-Jan-2000  
C:Accession: S32359  
R:Marchionni, M.A.; Goodearl, A.D.J.; Chen, M.S.; Bermingham-McDonogh, O.; Kirk, C.; Herles, I.; Davis, J.B.; Hsuan, J.J.; Totty, N.F.; Otsu, M.; McBurney, R.N.; Waterfield, M. Nature 362, 312-318, 1993  
A:Title: Glial growth factors are alternatively spliced erbB2 ligands expressed in the r  
A:Reference number: S32357; MUID:93205115  
A:Accession: S32359  
A:Molecule type: mRNA  
A:Residues: 1-241 <MAR>  
A:Cross-references: GB:U12259; NID:9289413; PIDN:AAA30540.1; PID:9289414  
C:Superfamily: unassigned EGF-related proteins; EGF homology  
F:182-221/Domain: EGF homology <EGF>

Query Match 37.7%; Score 104.5; DB 2; Length 241;  
Best Local Similarity 31.2%; Pred. No. 9.6e-05;  
Matches 15; Conservative 14; Mismatches 18; Indels 1; Gaps 1;  
OY 1 HFKPRDKDLAYCLNDGECFVETLTGSHKH-CRCKEGYGVRCDOFL 47  
DB 178 HLVKCAEKERTFCVNGGECFVWKDLSNPSRYLCKPCPNFTGDRCONY 225

RESULT 9

D43273  
heregulin precursor, splice form beta-3 - human  
N:Alternate names: glial growth factor HRC-beta-3; heregulin  
C:Species: Homo sapiens (man)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 11-Jan-2000  
C:Accession: D43273; S32358  
R:Holmes, W.E.; Sliwkowski, M.X.; Akita, R.M.; Henzel, W.J.; Lee, J.; Park, J.W.; Yan Science 256, 1205-1210, 1992  
A:Title: Identification of heregulin, a specific activator of p185(erbB2).  
A:Reference number: A43273; MUID:92271253  
A:Accession: D43273  
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual  
A:Molecule type: mRNA  
A:Residues: 1-241 <HOL>  
R:Marchionni, M.A.; Goodearl, A.D.J.; Chen, M.S.; Bermingham-McDonogh, O.; Kirk, C.; les, I.; Davis, J.B.; Hsuan, J.J.; Totty, N.F.; Otsu, M.; McBurney, R.N.; Waterfield, Nature 362, 312-318, 1993  
A:Title: Glial growth factors are alternatively spliced erbB2 ligands expressed in th  
A:Reference number: S32357; MUID:93205115  
A:Accession: S32358  
A:Molecule type: mRNA  
A:Residues: 1-241 <MAR>  
A:Cross-references: GB:U12261; NID:9292049; PIDN:AAB59358.1; PID:9292050  
C:Genetics:  
A:Gene: GDB:RGL; GGF  
A:Cross-references: GDB:132656; OMIM:142445  
A:Map position: 8p22-8p11  
C:Superfamily: unassigned EGF-related proteins; EGF homology  
C:Keywords: alternative splicing  
F:182-221/Domain: EGF homology <EGF>

Query Match 37.7%; Score 104.5; DB 2; Length 241;  
Best Local Similarity 31.2%; Pred. No. 9.6e-05;  
Matches 15; Conservative 14; Mismatches 18; Indels 1; Gaps 1;  
OY 1 HFKPRDKDLAYCLNDGECFVETLTGSHKH-CRCKEGYGVRCDOFL 47  
DB 178 HLVKCAEKERTFCVNGGECFVWKDLSNPSRYLCKPCPNFTGDRCONY 225

RESULT 10  
A56943  
sensory/motor neuron-derived factor - human  
C:Species: Homo sapiens (man)  
C>Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 11-Jan-2000  
C:Accession: A56943  
R:Ho, W.H.; Armanini, M.P.; Nuljens, A.; Phillips, H.S.; Osheroff, P.L. J. Biol. Chem. 270, 14523-14532, 1995  
A:Title: Sensory and motor neuron-derived factor. A novel heregulin variant highly ex  
A:Reference number: A56943; MUID:95301541  
A:Accession: A56943  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-296 <HOA>  
A:Cross-references: GB:U14827; NID:9862422; PIDN:AAC41764.1; PID:9862423  
C:Superfamily: unassigned EGF-related proteins; EGF homology  
F:237-276/Domain: EGF homology <EGF>

Query Match 37.7%; Score 104.5; DB 2; Length 296;  
Best Local Similarity 31.2%; Pred. No. 0.00011;  
Matches 15; Conservative 14; Mismatches 18; Indels 1; Gaps 1;  
OY 1 HFKPRDKDLAYCLNDGECFVETLTGSHKH-CRCKEGYGVRCDOFL 47  
DB 233 HLVKCAEKERTFCVNGGECFVWKDLSNPSRYLCKPCPNFTGDRCONY 280

RESULT 11  
S32357  
glial growth factor - human  
C:Species: Homo sapiens (man)

R:Holmes, W.E.; Slukowski, M.X.; Akita, R.W.; Henzel, W.T.; Lee, J.; Park, J.W.; Vancura

Accession: I61718  
Reference Number: A58210; M01D:94138863  
Status: Preliminary: translated from CB/FWR/CDDB1



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: September 13, 2002, 08:38:23 ; Search time 65.5 Seconds

(without alignments)  
79.702 Million cell updates/sec

Title: US-09-877-665-4

Sequence: 1 HRPCKDMLAVCLNDGECE.....SHKCKCKEYGVCRDQPL 47

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	277	100.0	47	20	AAW97622 Human neuroregulin r
2	277	100.0	157	20	AAW05451 Human heregulin-11
3	277	100.0	360	20	AAW97621 Human neuroregulin r
4	277	100.0	362	20	AAW97620 Mouse neuroregulin r
5	277	100.0	696	20	AAW97619 Human neuroregulin r
6	277	100.0	713	20	AAW97617 Mouse neuroregulin r
7	277	100.0	720	20	AAW05452 Human heregulin-11
8	277	100.0	720	20	AAW97618 Human neuroregulin r
9	116.5	42.1	52	17	AAW05182 Neu differentiation
10	116.5	42.1	52	21	AAW12602 Human NDF EGF-like
11	116.5	42.1	52	21	AAW69983 NDF/hergulin prot

12	113.5	41.0	52	17	AAW05184 Neu differentiation
13	113.5	41.0	63	15	AAW46918 EGF12. AAR46918;
14	113.5	41.0	63	15	AAW55659 EGF12. AAR55659;
15	113.5	41.0	63	16	AAW67250 Human epidermal 11
16	113.5	41.0	63	17	AAW09363 EGF12. Not specif
17	113.5	41.0	63	17	AAW96076 Epidermal growth f
18	113.5	41.0	63	17	AAW87461 Epidermal growth f
19	113.5	41.0	66	21	AAW36702 EGF-like domain of
20	113.5	41.0	83	15	AAW46922 EGF16. AAR46922;
21	113.5	41.0	83	15	AAW55663 EGF16. AAR55663;
22	113.5	41.0	83	16	AAW67254 Human epidermal 11
23	113.5	41.0	83	17	AAW09367 EGF10. Not specif
24	113.5	41.0	83	17	AAW36080 Epidermal growth f
25	113.5	41.0	83	17	AAW87465 Epidermal growth f
26	113.5	41.0	88	15	AAW46921 EGF15. AAR46921;
27	113.5	41.0	88	15	AAW55662 EGF15. AAR55662;
28	113.5	41.0	88	16	AAW67253 Human epidermal 11
29	113.5	41.0	88	17	AAW09366 Epidermal growth f
30	113.5	41.0	88	17	AAW36079 Epidermal growth f
31	113.5	41.0	88	17	AAW87464 Epidermal growth f
32	113.5	41.0	101	22	AAW67933 Human NRG1 protein
33	113.5	41.0	101	22	AAW67931 Human NRP-alpha3 C
34	113.5	41.0	125	22	AAW68564 Human NRG1 gene Ne
35	113.5	41.0	125	22	AAW67928 Human NRG1 gene Ne
36	113.5	41.0	125	22	AAW67966 Human NRG1 gene Ne
37	113.5	41.0	263	13	AAW28537 GGF2BP2. Bos tau
38	113.5	41.0	263	15	AAW46896 GGF2BP2. Bos tau
39	113.5	41.0	263	15	AAW55689 GGF2BP2. Bos tau
40	113.5	41.0	263	16	AAW67217 Putative bovine gl
41	113.5	41.0	263	17	AAW09360 Bovine neuroregulin
42	113.5	41.0	280	15	AAW46915 GGF2BP2. Bos tau
43	113.5	41.0	280	15	AAW55656 GGF2BP2. Bos tau
44	113.5	41.0	280	16	AAW67244 Bovine glial cell
45	113.5	41.0	280	17	AAW09369 Human neuroregulin G

#### ALIGNMENTS

RESULT 1

ID	AAW97622 standard; Protein: 47 AA.	ALIGNMENTS
XX	AAW97622;	
XX	10-MAY-1999 (first entry)	
DE	Human neuroregulin related ligand NRG3 EGF-like domain.	
KW	Neuroregulin related ligand; NRG3; hNRG3b1; human; ErbB4 receptor;	
KW	signal transduction; nervous system disorder; neurodegeneration;	
KW	neuropathy; therapy; diagnosis; epidermal growth factor; EGF;	
KW	immunoadhesin.	
XX	Homo sapiens.	
OS	MO9902681-AA.	
PN	21-JAN-1999.	
XX	30-JUN-1998; 98MO-US13411.	
PF	24-JUL-1997; 97US-0899437.	
PR	09-JUL-1997; 97US-0052019.	
XX	(GETH) GENENTECH INC.	
PA	Godowski PJ, Mark MR, Zhang D;	
XX	WPI; 1999-120882/10.	
XX	New isolated neuroregulin related ligand-3 - used to develop products	
PT	for treating nervous system disorders, e.g. stroke, ischemia.	

PT Infection, malignancy, Alzheimer's disease or Down's syndrome  
XX  
XX Claim 30; Page 64; 101pp; English.

CC This is the epidermal growth factor (EGF)-like domain of human  
CC neuroligin related ligand NR33 (see also AAM97618), a novel member of  
CC the EGF-like family of protein ligands that binds to the ErbB4  
CC receptor and activates ErbB4 receptor tyrosine phosphorylation.  
CC The EGF-1-like domain of NR33 is distinct from the EGF-like domains  
CC of NR31 and NR32. The invention provides human and murine  
CC polypeptides (see also AAM97617) that have at least 75% homology to  
CC the NR33 EGF-like domain, as well as expression vectors, host cells  
CC and methods for the recombinant production of novel NR33. The  
CC NR33 polypeptides and polynucleotides and can be used to enhance  
CC the survival, proliferation or differentiation of cells having the  
CC ErbB4 receptor in vivo and in vitro. They can be used to prevent  
CC or treat damage to a nerve or damage to other NR33-expressing or  
CC NR33-responsive cells, e.g. brain, heart, or kidney cells. In  
CC particular, they can be used to treat diseases which involve neural  
CC cell growth such as demyelination, or damage or loss of glial cells  
CC (e.g. multiple sclerosis). They can be used to treat patients whose  
CC nervous system has been damaged by e.g. trauma, surgery, stroke,  
CC ischemia, infection, metabolic disease, nutritional deficiency,  
CC malignancy, or toxic agents. NR33 can also be used to treat  
CC motor neuron disorders such as amyotrophic lateral sclerosis (Lou  
CC Gehrig's disease), Bell's palsy, conditions involving spinal  
CC muscular atrophy or paralysis, neurodegenerative disorders such as  
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple  
CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,  
CC and Meniere's disease. They can also be used to treat neuropathies  
CC associated with systemic disease including post-polio syndrome,  
CC hereditary neuropathies including Charcot-Marie-Tooth disease,  
CC Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's  
CC disease, metachromatic leukodystrophy, Fabry's disease and  
CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of  
CC skeletal muscle, such as muscular dystrophy or diseases caused by  
CC smooth muscle or smooth muscle wasting. The products can also be used  
CC for detection, diagnosis, for the production of transgenic or  
CC knockout animals or for drug screening. A claimed immunoglobulin  
CC sequence.

XX Sequence 47 AA;

QY Query Match 100.0%; Score 277; DB 20; Length 47;  
Best Local Similarity 100.0%; Pred. No. 2.5e-21;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 HFKPCRDRLAYCLNDGECFYETLTGSHKRCRCEGYGVRCDFL 47  
1 hfkpcrdkrlayclndgectfletltgshkrcrcegygvrcdfll 47

RESULT 2

AA05451  
ID AAY05451 standard; Protein; 157 AA.

XX AC AAY05451;

DT 06-JUL-1999 (first entry)

XX Human heregulin-like factor sequence.

KW Human heregulin-like factor; HLF; cell growth regulator; diagnosis;  
neural system disorder; cancer.

OS Homo sapiens.

XX MO9857968-A1.

PD 23-DEC-1998.

*X not good data*

PF 16-JUN-1998; 98MO-US12403.

XX 17-JUN-1997; 97US-0049942.

PA (HUMA-) HUMAN GENOME SCI INC.

XX (GEU) UNIV GEORGETOWN.

PI H1241 MM, King CR, Ruben SM, Young P;

XX WPI; 1999-095327/08.

DR N-PSDB; AAX36423.

PT New isolated heregulin-like factor - used to develop products for  
the diagnosis and treatment of disorders involving regulation of  
cell growth, particularly cancers

PS Claim 17; Page 86-87; 118pp; English.

CC This sequence is the human heregulin-like factor (HLF) of the  
CC invention. The HLF is involved in the regulation of cell growth.  
CC Detection of different levels of expression of the HLF gene can be used  
CC for the diagnosis of disorders, e.g. in the neural system. In  
CC particular, detection of different levels of HLF gene expression in cells  
CC or body fluid of an individual can be used for diagnosing cancer. The  
CC products can also be used in the treatment of disorders involving  
CC abnormal levels of HLF activity.

XX Sequence 157 AA;

QY Query Match 100.0%; Score 277; DB 20; Length 157;  
Best Local Similarity 100.0%; Pred. No. 7.6e-21;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 HFKPCRDRLAYCLNDGECFYETLTGSHKRCRCEGYGVRCDFL 47  
31 hfkpcrdkrlayclndgectfletltgshkrcrcegygvrcdfll 77

RESULT 3

AA097621  
ID AAM97621 standard; Protein; 360 AA.

XX AC AAM97621;

DT 10-MAY-1999 (first entry)

XX Human neuroligin related ligand NR33 extracellular domain.

KW Neuroligin related ligand; NR33; hNR33L; human; ErbB4 receptor;  
signal transduction; nervous system disorder; neurodegeneration;  
neuropathy; therapy; diagnosis.

OS Homo sapiens.

XX MO9902681-A1.

XX 21-JAN-1999.

XX 30-JUN-1998; 98MO-US13411.

XX 24-JUL-1997; 97US-0899437

XX 09-JUL-1997; 97US-0052049.

XX (GETH) GENENTECH INC.

XX Godowski P, Mark MR, Zhang D;

XX WPI; 1999-120882/10.

PT New isolated neuroligin related ligand-3 - used to develop products  
for treating nervous system disorders, e.g. stroke, ischemia,  
infection, malignancy, Alzheimer's disease or Down's syndrome

*X not good data*

PS Claim 5(a); Page 69-70; 101pp; English.

CC This is the extracellular domain (ECD, aa1-360 of human neuregulin  
CC related ligand NRG3 (see also AAM97618), a novel member of the  
CC epidermal growth factor (EGF)-like family of protein ligands. NRG3  
CC binds to the ErbB4 receptor, but not to the ErbB2 or ErbB3 receptor,  
CC activates ErbB4 receptor tyrosine phosphorylation. The invention  
CC provides human and murine polypeptides (see also AAM97617) that have  
CC at least 75% homology to the NRG3 ECD, as well as expression vectors,  
CC host cells and methods for the recombinant production of novel  
CC NRG3s. The NRG3 polypeptides and polynucleotides and can be used to  
CC enhance the survival, proliferation or differentiation of cells  
CC having the ErbB4 receptor in vivo and in vitro. They can be used to  
CC prevent or treat damage to a nerve or damage to other NRG3-expressing  
CC or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In  
CC particular, they can be used to treat diseases which involve neural  
CC cell growth such as demyelination, or damage or loss of glial cells  
CC (e.g. multiple sclerosis). They can be used to treat patients whose  
CC nervous system has been damaged by e.g. trauma, surgery, stroke,  
CC ischemia, infection, metabolic disease, nutritional deficiency,  
CC malignancy, or toxic agents. NRG3 can also be used to treat  
CC motor neuron disorders such as amyotrophic lateral sclerosis (Lou  
CC Gehrig's disease), Bell's palsy, conditions involving spinal  
CC muscular atrophy or paralysis, neurodegenerative disorders such as  
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple  
CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,  
CC and Meniere's disease. They can also be used to treat neuropathies  
CC associated with systemic disease including post-polio syndrome,  
CC hereditary neuropathies including Charcot-Marie-Tooth disease,  
CC Refsum's disease, abetalipoproteinemia, Tange's disease, Krabbe's  
CC disease, metachromatic leukodystrophy, Fabry's disease and  
CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of  
CC skeletal or smooth muscle wasting. The products can also be used  
CC for detection, diagnosis, for the production of transgenic or  
CC knockout animals or for drug screening.

Sequence 360 AA;

Query Match 100.0%; Score 277; DB 20; Length 360;  
Best Local Similarity 100.0%; Pred. No. 1.6e-20;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFRPCRDKLAYCLNDGCVFIEITLGSNHRCKRCKEGYGVRCDOFL 47  
DB 286 hfkpcrdklayclndgcvfietlgsnhckrcckegygvrddqfl 332

RESULT 4  
ID AAM97620 standard; Protein; 362 AA.

AAW97620;

10-MAY-1999 (first entry)

Mouse neuregulin related ligand NRG3 extracellular domain.

Neuregulin related ligand; NRG3; mouse; ErbB4 receptor;  
signal transduction; nervous system disorder; neurodegeneration;  
neuropathy; therapy; diagnosis.

Mus sp.

W09902681-A1.

21-JAN-1999.

30-JUN-1998; 98WO-US13411.

24-JUL-1997; 97US-0899437.

*Not good*

PR 09-JUL-1997; 97US-0052019.  
PA GENENTECH INC.  
PI Godowski PJ Mark MR, Zhang D;  
DR Ref. 1999-120882/10.

PS New isolated neuregulin related ligand-3 - used to develop products  
PT for treating nervous system disorders, e.g. stroke, ischemia,  
PT infection, malignancy, Alzheimer's disease or Down's syndrome  
XX Claim 5(a); Page 62-63; 101pp; English.

CC This is the extracellular domain (ECD, aa1-362) of murine neuregulin  
CC related ligand NRG3 (see also AAM97617), a novel member of the  
CC epidermal growth factor (EGF)-like family of protein ligands. NRG3  
CC binds to the ErbB4 receptor, but not to the ErbB2 or ErbB3 receptor,  
CC activates ErbB4 receptor tyrosine phosphorylation. The invention  
CC provides human and murine polypeptides (see also AAM97618) that have  
CC at least 75% homology to the NRG3 ECD, as well as expression vectors,  
CC host cells and methods for the recombinant production of novel  
CC NRG3s. The NRG3 polypeptides and polynucleotides and can be used to  
CC enhance the survival, proliferation or differentiation of cells  
CC having the ErbB4 receptor in vivo and in vitro. They can be used to  
CC prevent or treat damage to a nerve or damage to other NRG3-expressing  
CC or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In  
CC particular, they can be used to treat diseases which involve neural  
CC cell growth such as demyelination, or damage or loss of glial cells  
CC (e.g. multiple sclerosis). They can be used to treat patients whose  
CC nervous system has been damaged by e.g. trauma, surgery, stroke,  
CC ischemia, infection, metabolic disease, nutritional deficiency,  
CC malignancy, or toxic agents. NRG3 can also be used to treat  
CC motor neuron disorders such as amyotrophic lateral sclerosis (Lou  
CC Gehrig's disease), Bell's palsy, conditions involving spinal  
CC muscular atrophy or paralysis, neurodegenerative disorders such as  
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple  
CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,  
CC and Meniere's disease. They can also be used to treat neuropathies  
CC associated with systemic disease including post-polio syndrome,  
CC hereditary neuropathies including Charcot-Marie-Tooth disease,  
CC Refsum's disease, abetalipoproteinemia, Tange's disease, Krabbe's  
CC disease, metachromatic leukodystrophy, Fabry's disease and  
CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of  
CC skeletal or smooth muscle wasting. The products can also be used  
CC for detection, diagnosis, for the production of transgenic or  
CC knockout animals or for drug screening.

Sequence 362 AA;

Query Match 100.0%; Score 277; DB 20; Length 362;  
Best Local Similarity 100.0%; Pred. No. 1.6e-20;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFRPCRDKLAYCLNDGCVFIEITLGSNHRCKRCKEGYGVRCDOFL 47  
DB 288 hfkpcrdklayclndgcvfietlgsnhckrcckegygvrddqfl 334

RESULT 5  
ID AAM97619 standard; Protein; 696 AA.

AAW97619;

10-MAY-1999 (first entry)

Human neuregulin related ligand NRG3 (splice variant).

Neuregulin related ligand; NRG3; human; ErbB4 receptor;  
signal transduction; nervous system disorder; neurodegeneration;

neuropathy; therapy; diagnosis; splice variant.  
 Homo sapiens.  
 Key Location/Qualifiers  
 1..360 "extracellular domain, specifically claimed  
 in Claim 5(a)"  
 Region 66..91  
 Region /note= "hydrophobic region"  
 101..284  
 Region /note= "mucin-like Ser/Thr-rich region, contains  
 sites for O-linked glycosylation"  
 Domain 285..354  
 Domain /note= "EGF-like domain"  
 355..394  
 Domain /note= "transmembrane domain"  
 WO9902681-A1.  
 21-JAN-1999.  
 30-JUN-1998; 98WO-US13411.  
 24-JUL-1997; 97US-0899437.  
 09-JUL-1997; 97US-0052019.  
 (GENTH ) GENE TECH INC.  
 Godowski PJ, Mark MR, Zhang D;  
 WPI: 1999-120882/10.  
 N-PSDB; AAX06987.  
 New isolated neuregulin related ligand-3 - used to develop products  
 for treating nervous system disorders, e.g. stroke, ischemia,  
 infection, malignancy, Alzheimer's disease or Down's syndrome  
 Example 1; Page 78-81; 101pp; English.  
 This is the amino acid sequence of splice variant hNGR3B2 of human  
 neuregulin related ligand NRG3, a novel member of the epidermal  
 growth factor (EGF)-like family of protein ligands that binds to  
 the ErbB4 receptor, but not to the ErbB2 or ErbB3 receptor, and  
 which activates ErbB4 receptor tyrosine phosphorylation. The  
 sequence was deduced from the nucleotide sequence of a cDNA clone  
 (see AAX06989) from a foetal brain library. hNGR3B2 lacks amino  
 acids 529-552 of hNGR3B1 (see AAW97618) but retains the EGF-like  
 domain and is expected to exhibit biological activity. The invention  
 provides human and murine NRG3 polypeptides (see AAW97617), expression  
 vectors, host cells and methods for the recombinant production of  
 NRG3s. The NRG3 polypeptides and polynucleotides and can be used to  
 enhance the survival, proliferation or differentiation of cells  
 having the ErbB4 receptor in vivo and in vitro. They can be used to  
 prevent or treat damage to a nerve or damage to other NRG3-expressing  
 or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In  
 particular, they can be used to treat diseases which involve neural  
 cell growth such as demyelination, or damage or loss of glial cells  
 (e.g. multiple sclerosis). They can be used to treat patients whose  
 nervous system has been damaged by e.g. trauma, surgery, stroke,  
 ischemia, infection, metabolic disease, nutritional deficiency,  
 malignancy, or toxic agents. NRG3 can also be used to treat  
 motor neuron disorders such as amyotrophic lateral sclerosis (Lou  
 Gehrig's disease), Bell's palsy, conditions involving spinal  
 muscular atrophy or paralysis, neurodegenerative disorders such as  
 Alzheimer's disease, Huntington's chorea, Down's syndrome, multiple  
 sclerosis, Parkinson's disease, epilepsy, nerve deafness,  
 and Meniere's disease. They can also be used to treat neuropathies  
 associated with systemic disease including post-polio syndrome,  
 hereditary neuropathies including Charcot-Marie-Tooth disease,  
 Refsum's disease, abetalipoproteinemia, Tangle disease, Krabbe's  
 disease, metachromatic leukodystrophy, Fabry's disease and  
 Dejerine-Sottas syndrome, to treat disease of skeletal muscle of

smooth muscle, such as muscular dystrophy or diseases caused by  
 skeletal or smooth muscle wasting. The products can also be used  
 for detection, diagnosis, for the production of transgenic or  
 knockout animals or for drug screening.  
 Sequence 696 AA;  
 Query Match 100.0%; Score 277; DB 20; Length 696;  
 Best Local Similarity 100.0%; Pred. NO. 3e-20;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 HFKPCRDRLAYCLNDGCEFYLETITGSHKRCRKEGYQGVRCDFL 47  
 Db 286 hfkpcrdrlayclndgcefyletitgshkrcrckeygyrvrcdfli 332  
 RESULT 6  
 AAW97617  
 ID AAW97617 standard; Protein: 713 AA.  
 AC AAW97617;  
 DT 10-MAY-1999 (first entry)  
 DE Mouse neuregulin related ligand NRG3.  
 XX  
 KW Neuregulin related ligand; NRG3; mouse; ErbB4 receptor;  
 KW signal transduction; nervous system disorder; neurodegeneration;  
 KW neuropathy; therapy; diagnosis.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 1..362  
 FT /note= "extracellular domain, specifically claimed  
 in Claim 5(a)"  
 FT Region 66..91  
 FT /note= "hydrophobic region"  
 FT Region 105..286  
 FT /note= "mucin-like Ser/Thr-rich region, contains  
 sites for O-linked glycosylation"  
 FT Domain 287..334  
 FT /note= "EGF-like domain"  
 FT 363..385  
 FT Domain /note= "transmembrane domain"  
 WO9902681-A1.  
 21-JAN-1999.  
 30-JUN-1998; 98WO-US13411.  
 24-JUL-1997; 97US-0899437.  
 09-JUL-1997; 97US-0052019.  
 (GENTH ) GENE TECH INC.  
 Godowski PJ, Mark MR, Zhang D;  
 WPI: 1999-120882/10.  
 N-PSDB; AAX06987.  
 New isolated neuregulin related ligand-3 - used to develop products  
 for treating nervous system disorders, e.g. stroke, ischemia,  
 infection, malignancy, Alzheimer's disease or Down's syndrome  
 Claim 5(b); Page 59-62; 101pp; English.  
 This is the amino acid sequence of murine neuregulin related ligand  
 NRG3, a novel member of the epidermal growth factor (EGF)-like  
 family of protein ligands that binds to the ErbB4 receptor, but not  
 to the ErbB2 or ErbB3 receptor, and which activates ErbB4 receptor



tyrosine phosphorylation. The sequence was deduced from the nucleotide sequences of cDNA clones (see AAW06987) from a mouse brain library. The EGF-like domain of NRG3 is distinct from those of NRG1 or NRG2, and NRG3 displays receptor binding characteristics that are distinct from those of other neuroregulins. The invention provides human and murine NRG3 polypeptides (see also AAW97618), expression vectors, host cells and methods for the recombinant production of NRG3. The NRG3 polypeptides and polynucleotides and can be used to enhance the survival, proliferation or differentiation of cells having the ErbB4 receptor in vivo and in vitro. They can be used to prevent or treat damage to a nerve or damage to other NRG3-expressing or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In particular, they can be used to treat diseases which involve neural cell growth such as demyelination, or damage or loss of glial cells (e.g. multiple sclerosis). They can be used to treat patients whose nervous system has been damaged by e.g. trauma, surgery, stroke, malnutrition, infection, metabolic disease, nutritional deficiency, malignancy, or toxic agents. NRG3 can also be used to treat motor neuron disorders such as amyotrophic lateral sclerosis (Lou Gehrig's disease), Bell's palsy, conditions involving spinal muscular atrophy or paralysis, neurodegenerative disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's chorea, Down's syndrome, nerve deafness, and Meniere's disease. They can also be used to treat neuropathies associated with systemic disease including post-polio syndrome, hereditary neuropathies including Charcot-Marie-Tooth disease, Refsum's disease, abetalipoproteinemia, Tangle disease, Krabbe's disease, metachromatic leukodystrophy, Fabry's disease and Dejerine-Sottas syndrome, to treat disease or diseases caused by skeletal or smooth muscle wasting. The products can also be used for detection, diagnosis, for the production of transgenic or knockout animals or for drug screening.

Sequence 713 AA:

Query Match 100.0%; Score 277; DB 20; Length 713;  
Best Local Similarity 100.0%; Pred. No. 3.1e-20;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 HFKPCRDLDAYCLNDCECFYETLTGSHKRCRCKEGYGVRCDOFL 47  
286 hfkpcrdldayclndcecfyettltgshkrcrckegygvrddqfl 334

RESULT 7

AAW05452  
ID AAY05452 standard; Protein; 720 AA.

AC AAY05452;

DT 06-JUL-1999 (first entry)

DE Human heregulin-like factor sequence.

KW Human heregulin-like factor; HLF; cell growth regulator; diagnosis;  
neural system disorder; cancer.

OS Homo sapiens.

PN MO9857989-A1.

PD 23-DEC-1998.

PF 16-JUN-1998; 98WO-US12403.

PR 17-JUN-1997; 97US-0049942.

PA (HUMA-) HUMAN GENOME SCI INC.  
(GEOR) UNIV GEORGETOWN.

PI H1jazi MM, King CR, Ruben SM, Young P;

XX WPI: 1999-095327/08.  
DR New isolated heregulin-like factor - used to develop products for  
PT the diagnosis and treatment of disorders involving regulation of  
PT cell growth, particularly cancers  
XX  
PS Disclosure; Page 97-99; 118pp; English.  
XX  
CC This sequence is the human heregulin-like factor (HLF) of the  
CC invention. The HLF is involved in the regulation of cell growth.  
CC Detection of different levels of expression of the HLF gene can be used  
CC for the diagnosis of disorders, e.g. in the neural system. In  
CC particular, detection of different levels of HLF gene expression in cells  
CC or body fluid of an individual can be used for diagnosing cancer. The  
CC products can also be used in the treatment of disorders involving  
CC abnormal levels of HLF activity.  
XX  
SQ Sequence 720 AA:

Query Match 100.0%; Score 277; DB 20; Length 720;  
Best Local Similarity 100.0%; Pred. No. 3.1e-20;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 HFKPCRDLDAYCLNDCECFYETLTGSHKRCRCKEGYGVRCDOFL 47  
286 hfkpcrdldayclndcecfyettltgshkrcrckegygvrddqfl 332

RESULT 8

AAW97618  
ID AAW97618 standard; Protein; 720 AA.

AC AAW97618;

DT 10-MAY-1999 (first entry)

DE Human heregulin related ligand NRG3.

KW Neurogulin related ligand; NRG3; hNRG3B1; human; ErbB4 receptor;  
signal transduction; nervous system disorder; neurodegeneration;  
neuropathy; therapy; diagnosis.

OS Homo sapiens.

XX

FH Key

FT Domain

Location/Qualifiers  
1..360 "extracellular domain, specifically claimed  
in Claim 5(a)"

FT Region

66..91  
/note="hydrophobic region"

FT Region

101..284  
/note="mucin-like Ser/Thr-rich region, contains  
sites for O-linked glycosylation"

FT Domain

285..354  
/note="EGF-like domain"  
355..394  
/note="transmembrane domain"

PN WO9902681-A1.

PD 21-JAN-1999.

PF 30-JUN-1998; 98WO-US13411.

PR 24-JUL-1997; 97US-0899437.

PA (GEOR) GENENTECH INC.

PI Godowski PJ, Mark MR, Zhang D;

DR WPI: 1999-120882/10.  
 DR N-PSDB: AAM05182.  
 XX  
 PT New isolated neuregulin related ligand-3 - used to develop products  
 PT for treating nervous system disorders, e.g. stroke, ischemia,  
 PT infection, malignancy, Alzheimer's disease or Down's syndrome  
 XX  
 PS Claim 5(b); Page 66-69; 101pp; English.  
 XX  
 CC This is the amino acid sequence of human neuregulin related ligand  
 CC NR3, a novel member of the epidermal growth factor (EGF)-like  
 CC family of protein ligands that binds to the ErbB4 receptor, but not  
 CC to the ErbB2 or ErbB3 receptor, and which activates ErbB4 receptor  
 CC tyrosine phosphorylation. The sequence was deduced from the  
 CC nucleotide sequence of a cDNA clone (see AAM05182) from a fetal brain  
 CC library. The EGF-like domain of NR3 is distinct from those of NR1  
 CC or NR2, and NR3 displays receptor binding characteristics that are  
 CC distinct from those of other neuregulins. An alternatively spliced  
 CC form of human NR3 is provided in AAM97619. The invention provides  
 CC human and murine NR3 polypeptides (see also AAM97617), expression  
 CC vectors, host cells and methods for the recombinant production of  
 CC NR3s. The NR3 polypeptides and polynucleotides and can be used to  
 CC enhance the survival, proliferation or differentiation of cells  
 CC having the ErbB4 receptor in vivo and in vitro. They can be used to  
 CC prevent or treat damage to a nerve or damage to other NR3-expressing  
 CC or NR3-responsive cells, e.g. brain, heart, or kidney cells. In  
 CC particular, they can be used to treat diseases which involve neural  
 CC cell growth such as demyelination, or damage or loss of glial cells  
 CC (e.g. multiple sclerosis). They can be used to treat patients whose  
 CC nervous system has been damaged by e.g. trauma, surgery, stroke,  
 CC ischemia, infection, metabolic disease, nutritional deficiency,  
 CC malignancy, or toxic agents. NR3 can also be used to treat  
 CC motor neuron disorders such as amyotrophic lateral sclerosis (Lou  
 CC Gehrig's disease), Bell's palsy, conditions involving spinal  
 CC muscular atrophy or paralysis, neurodegenerative disorders such as  
 CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple  
 CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,  
 CC and Meniere's disease. They can also be used to treat neuropathies  
 CC associated with systemic disease including post-polio syndrome,  
 CC hereditary neuropathies including Charcot-Marie-Tooth disease,  
 CC Refsum's disease, abetalipoproteinemia, Tangle disease, Kriebel's  
 CC disease, metachromatic leukodystrophy, Fabry's disease and  
 CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of  
 CC smooth muscle, such as muscular dystrophy or diseases caused by  
 CC skeletal or smooth muscle wasting. The products can also be used  
 CC for detection, diagnosis, for the production of transgenic or  
 CC knockout animals or for drug screening.  
 CC  
 CC Sequence 720 AA;  
 SO  
 Query Match 100.0%; Score 277; DB 20; Length 720;  
 Best Local Similarity 100.0%; Pred. No. 3, 1e-20;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 HFKPRDRLAYCLNDGECFVETLTGSHKRCCKEGYGVRCDOFL 47  
 Db 266 hfkprdkldayclndgectvletltgshkrcckegygvrctdfl 332  
 RESULT 9  
 ID AAM05182 standard; peptide: 52 AA.  
 AC AAM05182;  
 DT 04-JUN-1997 (first entry)  
 XX Neu differentiation factor/hergulin-alpha/beta form EGF-like domain.  
 DE NDF; neu differentiation factor; heregulin; epidermal growth factor;  
 KW EGF; colon epithelial cell proliferation; Schwann cell; nerve;  
 DR damage; colitis; ulcer.

XX  
 OS Synthetic.  
 XX  
 PN WO9631599-A1.  
 XX  
 PD 10-OCT-1996.  
 XX  
 PF 27-MAR-1996; 96WO-US04262.  
 XX  
 PR 06-APR-1995; 95US-0417640.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX  
 PI Cranahan JF, Hara S, Lu HS, Mayer JP, Yoshinaga SK;  
 DR WPI: 1996-465022/46.  
 XX  
 PT Peptide(s) derived from neu differentiation factor/hergulin  
 PT proteins - specifically from epidermal growth factor-like domain,  
 PT stimulate proliferation of colon epithelial cells and Schwann cells  
 XX  
 PS Claim 1; Page 27; 37pp; English.  
 CC  
 CC The peptides AAM05182-W05185 are based on neu differentiation factor  
 CC (NDF)/heregulin alpha and beta form EGF-like domains in various  
 CC combinations. The peptides maintain the survival and proliferation of  
 CC Schwann cells and cause proliferation, growth and differentiation of  
 CC colon epithelial cells. Accordingly, they are useful to treat (in vitro  
 CC or in vivo) a disease or disorder of the colon (e.g. colitis or an  
 CC ulcer) or of the nervous system (e.g. nerve damage caused by trauma).  
 CC  
 CC Sequence 52 AA;  
 SO  
 Query Match 42.1%; Score 116.5; DB 17; Length 52;  
 Best Local Similarity 33.3%; Pred. No. 4, 7e-05;  
 Matches 16; Conservative 15; Mismatches 16; Indels 1; Gaps 1;  
 Oy 1 HFKPRDRLAYCLNDGECFVETLTGSHKRCCKEGYGVRCDOFL 47  
 Db 2 hlvkaeketkctvcgcefmwkdlnspnyrlckcpftgarcqnyv 49  
 RESULT 10  
 ID AAB12602 standard; Peptide: 52 AA.  
 AC AAB12602;  
 DT 09-NOV-2000 (first entry)  
 XX  
 DE Human NDF EGF-like domain derived peptide SEQ ID NO:1.  
 XX  
 KW Human; sensory epithelial cell; growth; stimulant; inner ear; EGF;  
 KW epithelial growth factor; NDF; heregulin; monoclonal antibody;  
 KW adult rat utricular epithelium.  
 XX  
 OS Homo sapiens.  
 PN US6080845-A.  
 XX  
 PD 27-JUN-2000.  
 XX  
 PF 28-JAN-1999; 99US-0238182.  
 XX  
 PR 05-AUG-1998; 98US-0129549.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX  
 PI Cranahan JF;  
 DR WPI: 2000-451228/39.  
 DR N-PSDB: AAM05182.



DE EGFL2.  
 XX  
 KW Glial growth factor; GGF; heregulin; mitogenesis;  
 KW Schwann cell; tumour; central nervous system; erbB2 receptor;  
 KW antiproliferative; epidermal growth factor; EGF.  
 XX  
 PN WO9403644-A.  
 XX  
 PD 17-FEB-1994.  
 XX  
 PF 10-AUG-1993; 93WO-US07491.  
 XX  
 PR 10-AUG-1992; 92US-0927337.  
 PR 25-SEP-1992; 92US-0951747.  
 PR 01-DEC-1992; 92US-0984085.  
 PR 29-JAN-1993; 93US-0011396.  
 XX  
 PA (CAMP-) CAMBRIDGE NEUROSCIENCE INC.  
 XX  
 PI Gwynne DI, Marchionni M, McBurney RN;  
 XX  
 DR WPI: 1994-065731/08.  
 DR N-PSDB; AAQ58324.  
 XX  
 PT Glial growth factor DNA encoding numerous polypeptide factors  
 PT used for inhibiting cell proliferation - for treating carcinoma  
 PT and nervous disorders  
 XX  
 PS Disclosure: Fig 40; 178pp; English.  
 XX  
 CC The GGF coding segments include regions with EGF-like homology.  
 CC These EGF-like domains can be required for the activation of  
 CC mitogenesis in the binding reaction between GGF ligands contg.  
 CC such domains and the erbB2 receptor. Pref. antiproliferative  
 CC factors are those which lack these EGF-like domains.  
 XX  
 SQ Sequence 63 AA;  
 OY  
 Query Match 41.0%; Score 113.5; DB 15; Length 63;  
 Best Local Similarity 34.8%; Pred. No. 0.00011;  
 Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;  
 Db 1 HFKPRDKLAVCLNDGCEFYIETITGSHKH-CRCKEYQGVRCQ 45  
 2 hlvxkaekctfcvngcefmwkdlnpsrylckcqpftgarcte 47  
 RESULT 14  
 AAR55659  
 ID AAR55659 standard; Protein; 63 AA.  
 XX  
 AC AAR55659;  
 XX  
 DT 28-JUL-1994 (first entry)  
 XX  
 DE EGFL2.  
 XX  
 KW Glial growth factor; GGF; heregulin; mitogenesis;  
 KW Schwann cell; tumour; central nervous system;  
 KW epidermal growth factor; EGF.  
 XX  
 PN WO9400140-A.  
 XX  
 PD 06-JAN-1994..  
 XX  
 PF 29-JUN-1993; 93WO-US06228.  
 XX  
 PR 30-JUN-1992; 92US-0907138.  
 PR 03-SEP-1992; 92US-0940389.  
 PR 23-OCT-1992; 92US-0965173.  
 PR 24-MAR-1993; 93US-0036555.  
 XX

PA (CAMP-) CAMBRIDGE NEUROSCIENCE.  
 PA (LUDM-) LUDMIG INST CANCER RES.  
 XX  
 PI Chen MS, Goodearl A, Hiles I, Marchionni M, Minghetti L;  
 PI Stroobant P, Waterfield M;  
 XX  
 DR WPI: 1994-025882/03.  
 DR N-PSDB; AAQ62843.  
 XX  
 PT Glial mitogenic polypeptide factors - useful for stimulating  
 PT glial cell mitogenesis and treating glial cell tumours  
 XX  
 PS Claim 53; Fig 39; 178pp; English.  
 XX  
 CC EGFL1, EGFL2, EGFL3, EGFL4, EGFL5 and EGFL6 are used  
 CC for the stimulation of glial cell mitogenesis in vivo  
 CC and in vitro.  
 XX  
 SQ Sequence 63 AA;  
 OY  
 Query Match 41.0%; Score 113.5; DB 15; Length 63;  
 Best Local Similarity 34.8%; Pred. No. 0.00011;  
 Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;  
 Db 1 HFKPRDKLAVCLNDGCEFYIETITGSHKH-CRCKEYQGVRCQ 45  
 2 hlvxkaekctfcvngcefmwkdlnpsrylckcqpftgarcte 47  
 RESULT 15  
 AAR67250  
 ID AAR67250 standard; Protein; 63 AA.  
 XX  
 AC AAR67250;  
 XX  
 DT 15-AUG-1995 (first entry)  
 XX  
 DE Human epidermal like growth factor 2 (EGFL2).  
 XX  
 KW Epidermal like growth factor 2; mammalian muscle cell treatment;  
 KW skeletal; cardiac; smooth; acetylcholine receptor deficiency;  
 KW EGFL2.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9426298-A.  
 XX  
 PD 24-NOV-1994.  
 XX  
 PF 06-MAY-1994; 94WO-US05083.  
 XX  
 PR 06-MAY-1993; 93US-0059022.  
 PR 08-MAR-1994; 94US-0209204.  
 XX  
 PA (CAMP-) CAMBRIDGE NEUROSCIENCE.  
 XX  
 PI Gwynne DI, Marchionni M, Sklar R;  
 PI WPI: 1995-006353/01.  
 DR N-PSDB; AAQ74915.  
 XX  
 PT Treating mammalian muscle diseases and disorders - by admin. of  
 PT GGF2 and other specified polypeptide(s) which bind the p185erbB2  
 PT receptor.  
 XX  
 PS Claim 34; Pages 148-149; 241pp; English.  
 XX  
 CC AAQ7915 encodes AAR67250 human epidermal like growth factor 2 (EGFL2).  
 CC The glial cell mitogenic activity of EGFL2 can be used to treat a  
 CC variety of mammalian skeletal, cardiac and smooth muscle diseases,  
 CC including acetylcholine receptor deficiency.  
 XX

